NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1D7 USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND OTHER CANCERS

This application is a continuation-in-part of U.S. patent application 10/280,340 filed 25 October 2002 and a continuation-in-part of U.S. patent application 10/277,292 filed 21October 2002, both of which are continuations of U.S. Patent Application No. 09/935,430 filed August 22, 2001, which claimed the priority benefit of United States Provisional Patent Applications 60/227,098, filed August 22, 2000, and 60/282,739, filed April 10, 2001. This application also claims priority benefit of United States Provisional Patent Application 60/446,633, filed February 10, 2003. The entire contents of all of these applications are fully incorporated herein by reference.

FIELD OF THE INVENTION

The invention described herein relates to novel nucleic acid sequences and thier encoded proteins, referred to as 158P1D7 and variants thereof, and to diagnostic and therapeutic methods and compositions useful in the management of various cancers that express 158P1D7 and variants thereof.

BACKGROUND OF THE INVENTION

Cancer is the second leading cause of human death next to coronary disease. Worldwide, millions of people die from cancer every year. In the United States alone, as reported by the American Cancer Society, cancer causes the death of well over a half-million people annually, with over 1.2 million new cases diagnosed per year. While deaths from heart disease have been declining significantly, those resulting from cancer generally are on the rise. In the early part of the next century, cancer is predicted to become the leading cause of death.

Of all new cases of cancer in the United States, bladder cancer represents approximately 5 percent in men (fifth most common neoplasm) and 3 percent in women (eighth most common neoplasm). The incidence is increasing slowly, concurrent with an increasing older population. In 1998, there was an estimated 54,500 cases, including 39,500 in men and 15,000 in women. The age-adjusted incidence in the United States is 32 per 100,000 for men and 8 per 100,000 in women. The historic male/female ratio of 3:1 may be decreasing related to smoking patterns in women. There were an estimated 11,000 deaths from bladder cancer in 1998 (7,800 in men and 3,900 in women). Bladder cancer incidence and mortality strongly increase with age and will be an increasing problem as the population becomes more elderly.

Bladder cancers comprise a heterogeneous group of diseases. The main determinants of disease control and survival are histology and extent of disease. The main codes for these factors include pathology classification, the International Classification of Diseases-Oncology (ICDO), and staging classification of extent of disease, the TNM classification.(Table XXI). For a general discussion of bladder and other urogenital cancers, see, e.g., Volgelzang, et al, Eds. Comprehensive Textbook of Genitourinary Oncology, (Williams & Wilkins, Baltimore 1996), in particular pages 295-556.

Three primary types of tumors have been reported in the bladder. The most common type of bladder cancer is Transitional cell carcinoma (TCC); this accounts for about 90% of all bladder cancers. The second form of bladder cancer is squamous cell carcinoma, which accounts for about 8% of all bladder cancers where schistosomiasis is not endemic, and approximately 75% of bladder carcinomas where schistosomiasis is endemic. Squamous cell carcinomas tend to invade deeper layers of the bladder. The third type of bladder cancer is adenocarcinoma, which account for 1%-2% of bladder cancers; these are primarily invasive forms of cancer.

Bladder cancer is commonly detected and diagnosed using cytoscopy and urine cytology. However these methods demonstrate poor sensitivity. Relatively more reliable methods of detection currently used in the clinic include the bladder tumor antigen (BTA) stat test, NMP22 protein assay, telomerase expression and hyaluronic acid and hyaluronidase (HA-

HAase) urine test. The advantage of using such markers in the diagnosis of bladder cancer is their relative high sensitivity in earlier tumor stages compared to standard cytology.

For example, the BTA stat test has 60-80% sensitivity and 50-70% specificity for bladder cancer , while the HA-HAase urine test shows 90-92% sensitivity and 80-84% specificity for bladder cancer (J Urol 2001 165:1067). In general, sensitivity for stage Ta tumors was 81% for nuclear matrix protein (NMP22), 70% for telomerase, 32% for bladder tumor antigen (BTA) and 26% for cytology (J Urol 2001 166:470; J Urol 1999, 161:810). Although the telomeric repeat assay which measures telomerase activity is relatively sensitive, instability of telomerase in urine presently renders this detection method unreliable.

Most bladder cancers recur in the bladder. Generally, bladder cancer is managed with a combination of transurethral resection of the bladder (TUR) and intravesical chemotherapy or immunotherapy. The multifocal and recurrent nature of bladder cancer points out the limitations of TUR. Most muscle-invasive cancers are not cured by TUR alone. Radical cystectomy and urinary diversion is the most effective means to eliminate the cancer but carry an undeniable impact on urinary and sexual function.

Intravesical bacilli Calmette-Guerin (BCG) is a common and efficacious immunotherapeutic agent used in the treatment of bladder cancer. BCG is also used as a prophylactic agent to prevent recurrence of bladder cancer. However, 30% of patients fail to respond to BCG therapy and go on to develop invasive and metastatic disease (Catalona et al. J Urol 1987, 137:220-224). BCG-related side effects have been frequently observed such as drug-induced cystitis, risk of bacterial infection, and hematuria, amongst others. Other alternative immunotherapies have been used for the treatment of bladder cancer, such as KLH (Flamm et al. Urologe 1994; 33:138-143) interferons (Bazarbashi et al. J Surg Oncol. 2000; 74:181-4), and MAGE-3 peptide loaded dendritic cells (Nishiyama et al. Clin Cancer Res 2001; 7:23-31). All these approaches are still experimental (Zlotta et al. Eur Urol 2000;37 Suppl 3:10-15). There continues to be a significant need for diagnostic and treatment modalities that are beneficial for bladder cancer patients. Furthermore, from a worldwide standpoint, several cancers stand out as the leading killers. In particular, carcinomas of the lung, prostate, breast, colon, pancreas, and ovary are primary causes of cancer death. These and virtually all other carcinomas share a common lethal feature. With very few exceptions, metastatic disease from a carcinoma is fatal. Moreover, even for those cancer patients who initially survive their primary cancers, their lives are dramatically altered. Many cancer patients experience strong anxieties driven by the awareness of the potential for recurrence or treatment failure. Many cancer patients experience physical debilitations following treatment. Furthermore, many cancer patients experience a recurrence.

Prostate cancer is the fourth most prevalent cancer in men worldwide. In North America and Northern Europe, it is by far the most common cancer in males and is the second leading cause of cancer death in men. In the United States alone, well over 30,000 men die annually of this disease, second only to lung cancer. Despite the magnitude of these figures, there is still no effective treatment for metastatic prostate cancer. Surgical prostatectomy, radiation therapy, hormone ablation therapy, surgical castration and chemotherapy continue to be the main treatment modalities. Unfortunately, these treatments are ineffective for many and are often associated with undesirable consequences.

On the diagnostic front, the lack of a prostate tumor marker that can accurately detect early-stage, localized tumors remains a significant limitation in the diagnosis and management of this disease. Although the serum prostate specific antigen (PSA) assay has been a very useful tool, however its specificity and general utility is widely regarded as lacking in several important respects. While previously identified markers such as PSA, PSM, PCTA and PSCA have facilitated efforts to diagnose and treat prostate cancer, there is need for the identification of additional markers and therapeutic targets for prostate and related cancers in order to further improve diagnosis and therapy.

Renal cell carcinoma (RCC) accounts for approximately 3 percent of adult malignancies. Once adenomas reach a diameter of 2 to 3 cm, malignant potential exists. In the adult, the two principal malignant renal tumors are renal cell

adenocarcinoma and transitional cell carcinoma of the renal pelvis or ureter. The incidence of renal cell adenocarcinoma is estimated at more than 29,000 cases in the United States, and more than 11,600 patients died of this disease in 1998. Transitional cell carcinoma is less frequent, with an incidence of approximately 500 cases per year in the United States.

Surgery has been the primary therapy for renal cell adenocarcinoma for many decades. Until recently, metastatic disease has been refractory to any systemic therapy. With recent developments in systemic therapies, particularly immunotherapies, metastatic renal cell carcinoma may be approached aggressively in appropriate patients with a possibility of durable responses. Nevertheless, there is a remaining need for effective therapies for these patients.

An estimated 130,200 cases of colorectal cancer occurred in 2000 in the United States, including 93,800 cases of colon cancer and 36,400 of rectal cancer. Colorectal cancers are the third most common cancers in men and women. Incidence rates declined significantly during 1992-1996 (-2.1% per year). Research suggests that these declines have been due to increased screening and polyp removal, preventing progression of polyps to invasive cancers. There were an estimated 56,300 deaths (47,700 from colon cancer, 8,600 from rectal cancer) in 2000, accounting for about 11% of all U.S. cancer deaths.

At present, surgery is the most common form of therapy for colorectal cancer, and for cancers that have not spread, it is frequently curative. Chemotherapy, or chemotherapy plus radiation is given before or after surgery to most patients whose cancer has deeply perforated the bowel wall or has spread to the lymph nodes. A permanent colostomy (creation of an abdominal opening for elimination of body wastes) is occasionally needed for colon cancer and is infrequently required for rectal cancer. There continues to be a need for effective diagnostic and treatment modalities for colorectal cancer.

There were an estimated 164,100 new cases of lung and bronchial cancer in 2000, accounting for 14% of all U.S. cancer diagnoses. The incidence rate of lung and bronchial cancer is declining significantly in men, from a high of 86.5 per 100,000 in 1984 to 70.0 in 1996. In the 1990s, the rate of increase among women began to slow. In 1996, the incidence rate in women was 42.3 per 100,000.

Lung and bronchial cancer caused an estimated 156,900 deaths in 2000, accounting for 28% of all cancer deaths. During 1992–1996, mortality from lung cancer declined significantly among men (-1.7% per year) while rates for women were still significantly increasing (0.9% per year). Since 1987, more women have died each year of lung cancer than breast cancer, which, for over 40 years, was the major cause of cancer death in women. Decreasing lung cancer incidence and mortality rates most likely resulted from decreased smoking rates over the previous 30 years; however, decreasing smoking patterns among women lag behind those of men. Of concern, although the declines in adult tobacco use have slowed, tobacco use in youth is increasing again.

Treatment options for lung and bronchial cancer are determined by the type and stage of the cancer and include surgery, radiation therapy, and chemotherapy. For many localized cancers, surgery is usually the treatment of choice. Because the disease has usually spread by the time it is discovered, radiation therapy and chemotherapy are often needed in combination with surgery. Chemotherapy alone or combined with radiation is the treatment of choice for small cell lung cancer; on this regimen, a large percentage of patients experience remission, which in some cases is long lasting. There is however, an ongoing need for effective treatment and diagnostic approaches for lunch and bronchial cancers.

An estimated 182,800 new invasive cases of breast cancer were expected to have occurred among women in the United States during 2000. Additionally, about 1,400 new cases of breast cancer were expected to be diagnosed in men in 2000. After increasing about 4% per year in the 1980s, breast cancer incidence rates in women have leveled off in the 1990s to about 110.6 cases per 100,000.

In the U.S. alone, there were an estimated 41,200 deaths (40,800 women, 400 men) in 2000 due to breast cancer. Breast cancer ranks second among cancer deaths in women. According to the most recent data, mortality rates declined

significantly during 1992–1996 with the largest decreases in younger women, both white and black. These decreases were probably the result of earlier detection and improved treatment.

Taking into account the medical circumstances and the patient's preferences, treatment of breast cancer may involve lumpectomy (local removal of the tumor) and removal of the lymph nodes under the arm; mastectomy (surgical removal of the breast) and removal of the lymph nodes under the arm; radiation therapy; chemotherapy; or hormone therapy. Often, two or more methods are used in combination. Numerous studies have shown that, for early stage disease, long-term survival rates after lumpectomy plus radiotherapy are similar to survival rates after modified radical mastectomy. Significant advances in reconstruction techniques provide several options for breast reconstruction after mastectomy. Recently, such reconstruction has been done at the same time as the mastectomy.

Local excision of ductal carcinoma in situ (DCIS) with adequate amounts of surrounding normal breast tissue may prevent the local recurrence of the DCIS. Radiation to the breast and/or tamoxifen may reduce the chance of DCIS occurring in the remaining breast tissue. This is important because DCIS, if left untreated, may develop into invasive breast cancer. Nevertheless, there are serious side effects or sequelae to these treatments. There is, therefore, a need for efficacious breast cancer treatments.

There were an estimated 23,100 new cases of ovarian cancer in the United States in 2000. It accounts for 4% of all cancers among women and ranks second among gynecologic cancers. During 1992–1996, ovarian cancer incidence rates were significantly declining. Consequent to ovarian cancer, there were an estimated 14,000 deaths in 2000. Ovarian cancer causes more deaths than any other cancer of the female reproductive system.

Surgery, radiation therapy, and chemotherapy are treatment options for ovarian cancer. Surgery usually includes the removal of one or both ovaries, the fallopian tubes (salpingo-oophorectomy), and the uterus (hysterectomy). In some very early tumors, only the involved ovary will be removed, especially in young women who wish to have children. In advanced disease, an attempt is made to remove all intra-abdominal disease to enhance the effect of chemotherapy. There continues to be an important need for effective treatment options for ovarian cancer.

There were an estimated 28,300 new cases of pancreatic cancer in the United States in 2000. Over the past 20 years, rates of pancreatic cancer have declined in men. Rates among women have remained approximately constant but may be beginning to decline. Pancreatic cancer caused an estimated 28,200 deaths in 2000 in the United States. Over the past 20 years, there has been a slight but significant decrease in mortality rates among men (about –0.9% per year) while rates have increased slightly among women.

Surgery, radiation therapy, and chemotherapy are treatment options for pancreatic cancer. These treatment options can extend survival and/or relieve symptoms in many patients but are not likely to produce a cure for most. There is a significant need for additional therapeutic and diagnostic options for pancreatic cancer.

SUMMARY OF THE INVENTION

The present invention relates to a novel nucleic acid sequence and its encoded polypeptide, designated 158P1D7. As used herein, "158P1D7" may refer to the novel polynucleotides or polypeptides or variants thereof or both of the disclosed invention.

Nucleic acids encoding 158P1D7 are over-expressed in the cancer(s) listed in Table I. Northern blot expression analysis of 158P1D7 expression in normal tissues shows a restricted expression pattern in adult tissues. The nucleotide (Figure 2) and amino acid (Figure 2, and Figure 3) sequences of 158P1D7 are provided. The tissue-related profile of 158P1D7 in normal adult tissues, combined with the over-expression observed in bladder tumors, shows that 158P1D7 is aberrantly over-expressed in at least some cancers. Thus, 158P1D7 nucleic acids and polypeptides serve as a useful diagnostic agent (or indicator) and/or therapeutic target for cancers of the tissues, such as those listed in Table I.

The invention provides polynucleotides corresponding or complementary to all or part of the 158P1D7 nucleic acids, mRNAs, and/or coding sequences, preferably in isolated form, including polynucleotides encoding 158P1D7-related proteins and fragments of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or more than 25 contiguous amino acids; at least about 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100 or more than 100 contiguous amino acids of a 158P1D7-related protein, as well as the peptides/proteins themselves; DNA, RNA, DNA/RNA hybrids, and related molecules (such as PNAs), polynucleotides or oligonucleotides complementary or having at least a 90% homology to 158P1D7 nucleic acid sequences or mRNA sequences or parts thereof, and polynucleotides or oligonucleotides that hybridize to the 158P1D7 genes, mRNAs, or to 158P1D7-encoding polynucleotides. Also provided are means for isolating cDNAs and the gene(s) encoding 158P1D7. Recombinant DNA molecules containing 158P1D7 polynucleotides, cells transformed or transduced with such molecules, and host-vector systems for the expression of 158P1D7 gene products are also provided. The invention further provides antibodies that bind to 158P1D7 proteins and polypeptide fragments thereof, including polyclonal and monoclonal antibodies, murine and other mammalian antibodies, chimeric antibodies, humanized and fully human antibodies, and antibodies labeled with a detectable marker. The invention also comprises T cell clones that recognize an epitope of 158P1D7 in the context of a particular HLA molecule.

The invention further provides methods for detecting the presence, amount, and status of 158P1D7 polynucleotides and proteins in various biological samples, as well as methods for identifying cells that express 158P1D7 polynucleotides and polypeptides. A typical embodiment of this invention provides methods for monitoring 158P1D7 polynucleotides and polypeptides in a tissue or hematology sample having or suspected of having some form of growth dysregulation such as cancer.

Note that to determine the starting position of any peptide set forth in Tables V-XVIII and XXII to XLIX (collectively HLA Peptide Tables) respective to its parental protein, e.g., variant 1, variant 2, etc., reference is made to three factors: the particular variant, the length of the peptide in an HLA Peptide Table, and the Search Peptides in Table VII. Generally, a unique Search Peptide is used to obtain HLA peptides of a particular for a particular variant. The position of each Search Peptide relative to its respective parent molecule is listed in Table 55. Accordingly, if a Search Peptide begins at position "X", one must add the value "X - 1" to each position in Tables V-XVIII and XXII to XLIX to obtain the actual position of the HLA peptides in their parental molecule. For example, if a particular Search Peptide begins at position 150 of its parental molecule, one must add 150 - 1, i.e., 149 to each HLA peptide amino acid position to calculate the position of that amino acid in the parent molecule.

The invention further provides various immunogenic or therapeutic compositions and strategies for treating cancers that express 158P1D7 such as bladder cancers, including therapies aimed at inhibiting the transcription, translation, processing or function of 158P1D7 as well as cancer vaccines.

The invention further provides a method of generating a mammalian immune response directed to a protein of Figure 2, where the method comprises exposing cells of the mammal's immune system to a portion of a) a 158P1D7-related protein and/or b) a nucleotide sequence that encodes said protein, whereby an immune response is generated to said protein. The 158P1D7-related protein can comprise at least one T cell or at least one B cell epitope; and, upon contacting the epitope with a mammalian immune system T cell or B cell respectively, the T cell or B cell is activated. The immune system cell is a B cell, a cytotoxic T cell (CTL), and/or a helper T cell (HTL). When the immune system cell is a B cell, the activated B cell generates antibodies that specifically bind to the 158P1D7-related protein. When the immune system cell is a T cell that is a cytotoxic T cell (CTL), the activated CTL kills an autologous cell that expresses the 158P1D7-related protein. When the immune system cell is a T cell that is a helper T cell (HTL), the activated HTL secretes cytokines that facilitate the cytotoxic activity of a cytotoxic T cell (CTL) or the antibody-producing activity of a B cell.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. 158P1D7 SSH nucleic acid sequence. The 158P1D7 SSH sequence contains 231 bp.

Figure 2. A) The cDNA and amino acid sequence of 158P1D7 variant 1 (also called "158P1D7 v.1" or "158P1D7 variant 1") is shown in Figure 2A. The start methionine is underlined. The open reading frame extends from nucleic acid 23-2548 including the stop codon.

B) The cDNA and amino acid sequence of 158P1D7 variant 2 (also called "158P1D7 v.2") is shown in Figure 2B. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 23-2548 including the stop codon.

C) The cDNA and amino acid sequence of 158P1D7 variant 3 (also called "158P1D7 v.3") is shown in Figure 2C. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 23-2221 including the stop codon.

D) The cDNA and amino acid sequence of 158P1D7 variant 4 (also called "158P1D7 v.4") is shown in Figure 2D. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 23-1210 including the stop codon.

E) The cDNA and amino acid sequence of 158P1D7 variant 5 (also called "158P1D7 v.5") is shown in Figure 2E. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 480-3005 including the stop codon.

F) The cDNA and amino acid sequence of 158P1D7 variant 6 (also called "158P1D7 v.6") is shown in Figure 2F. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 23-1612 including the stop codon.

Figure 3.

A) The amino acid sequence of 158P1D7 v.1 is shown in Figure 3A; it has 841 amino acids.

B) The amino acid sequence of 158P1D7 v.3 is shown in Figure 3B; it has 732 amino acids.

C) The amino acid sequence of 158P1D7 v.4 is shown in Figure 3C; it has 395 amino acids.

D) The amino acid sequence of 158P1D7 v.6 is shown in Figure 3D; it has 529 amino acids.

As used herein, a reference to 158P1D7 includes all variants thereof, including those shown in Figures 2, 3, 10, 11, and 12 unless the context clearly indicates otherwise.

Figure 4. Alignment BLAST homology of 158P1D7 v.1 amino acid to hypothetical protein FLJ22774.

Figure 5. Figure 5a: Amino acid sequence alignment of 158P1D7 with human protein. **Figure 5b:** Amino acid sequence alignment of 158P1D7 with human protein similar to IGFALS.

Figure 6. Expression of 158P1D7 by RT-PCR. First strand cDNA was prepared from vital pool 1 (VP1: liver, lung and kidney), vital pool 2 (VP2, pancreas, colon and stomach), prostate xenograft pool (LAPC-4AD, LAPC-4AI, LAPC-9AD, LAPC-9AI), prostate cancer pool, bladder cancer pool , colon cancer pool , lung cancer pool, ovary cancer pool, breast cancer pool, and metastasis pool. Normalization was performed by PCR using primers to actin and GAPDH. Semi-quantitative PCR, using primers to 158P1D7, was performed at 30 cycles of amplification. Strong expression of 158P1D7 is observed in bladder cancer pool and breast cancer pool. Lower levels of expression are observed in VP1, VP2, xenograft pool, prostate cancer pool, colon cancer pool , lung cancer pool, ovary cancer pool, and metastasis pool.

Figure 7. Expression of 158P1D7 in normal human tissues. Two multiple tissue northern blots, with 2 μg of mRNA/lane, were probed with the 158P1D7 fragment. Size standards in kilobases (kb) are indicated on the side. The results show expression of 158P1D7 in prostate, liver, placenta, heart and, to lower levels, in small intestine and colon.

Figure 8. Expression of 158P1D7 in bladder cancer patient specimens. Figure 8A. RNA was extracted from the bladder cancer cell lines (CL), normal bladder (N), bladder tumors (T) and matched normal adjacent tissue (N_{AT}) isolated

from bladder cancer patients. Northern blots with 10 μg of total RNA/lane were probed with the 158P1D7 fragment. Size standards in kilobases (kb) are indicated on the side. The results show expression of 158P1D7 in 1 of 3 bladder cancer cell lines. In patient specimens, 158P1D7 expression is detected in 4 of 6 tumors tested. **Figure 8B.** In another study, 158P1D7 expression is detected in all patient tumors tested (8B). The expression observed in normal adjacent tissues (isolated from diseased tissues) but not in normal tissue, isolated from healthy donors, may indicate that these tissues are not fully normal and that 158P1D7 may be expressed in early stage tumors.

Figure 9. Expression of 158P1D7 in lung cancer patient specimens. RNA was extracted from lung cancer cell lines (CL), lung tumors (T), and their normal adjacent tissues (N_{AT}) isolated from lung cancer patients. Northern blot with 10 μg of total RNA/lane was probed with the 158P1D7 fragment. Size standards in kilobases (kb) are indicated on the side. The results show expression of 158P1D7 in 1 of 3 lung cancer cell lines and in all 3 lung tumors tested, but not in normal lung tissues.

Figure 10. Expression of 158P1D7 in breast cancer patient specimens. RNA was extracted from breast cancer cell lines (CL), normal breast (N), and breast tumors (T) isolated from breast cancer patients. Northern blot with 10 μg of total RNA/lane was probed with the 158P1D7 fragment: Size standards in kilobases (kb) are indicated on the side. The results show expression of 158P1D7 in 2 of 3 breast cancer cell lines and in 2 breast tumors, but not in normal breast tissue.

Figure 11. Figures 11(a) – (d): Hydrophilicity amino acid profile of 158P1D7 v.1, v.3, v.4, and v.6 determined by computer algorithm sequence analysis using the method of Hopp and Woods (Hopp T.P., Woods K.R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828) accessed on the Protscale website located on the World Wide Web at (URL: expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

Figure 12. Figures 12(a)-(d): Hydropathicity amino acid profile of 158P1D7 v.1, v.3, v.4, and v.6 determined by computer algorithm sequence analysis using the method of Kyte and Doolittle (Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132) accessed on the ProtScale website located on the World Wide Web at (URL: expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

Figure 13. Figures 13(a)-(d): Percent accessible residues amino acid profile of 158P1D7 v.1, v.3, v.4, and v.6 determined by computer algorithm sequence analysis using the method of Janin (Janin J., 1979 Nature 277:491-492) accessed on the ProtScale website located on the World Wide Web at (URL: expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

Figure 14. Figures 14(a)-(d): Average flexibility amino acid profile of 158P1D7 v.1, v.3, v.4, and v.6 determined by computer algorithm sequence analysis using the method of Bhaskaran and Ponnuswamy (Bhaskaran R., and Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255) accessed on the ProtScale website located on the World Wide Web at (URL: expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

Figure 15. Figures 15(a)-(d): Beta-turn amino acid profile of 158P1D7 v.1, v.3, v.4, and v.6 determined by computer algorithm sequence analysis using the method of Deleage and Roux (Deleage, G., Roux B. 1987 Protein Engineering 1:289-294) accessed on the ProtScale website located on the World Wide Web at (URL: expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

Figure 16. Figures 16(A)-(D): Secondary structure and transmembrane domains prediction for 158P1D7 protein variants. The secondary structures of 158P1D7 protein variants 1 (SEQ ID NO: 104), v.3 (SEQ ID NO: 105), v.4 (SEQ ID NO: 106), and v.6 (SEQ ID NO: 107), respectively, were predicted using the HNN - Hierarchical Neural Network method (NPS@: Network Protein Sequence Analysis TIBS 2000 March Vol. 25, No 3 [291]:147-150 Combet C., Blanchet C., Geourjon C. and Deléage G., http://pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html), accessed from the ExPasy

molecular biology server located on the World Wide Web at (.expasy.ch/tools/). This method predicts the presence and location of alpha helices, extended strands, and random coils from the primary protein sequence. The percent of the protein variant in a given secondary structure is also listed. **Figures 16E, 16G, 16I, and 16K:** Schematic representation of the probability of existence of transmembrane regions of 158P1D7 protein variants 1, 3, 4, and 6, respectively, based on the TMpred algorithm of Hofmann and Stoffel which utilizes TMBASE (K. Hofmann, W. Stoffel. TMBASE - A database of membrane spanning protein segments Biol. Chem. Hoppe-Seyler 374:166, 1993). **Figures 16F, 16H, 16J, and 16L:** Schematic representation of the probability of the existence of transmembrane regions of 158P1D7 protein variants 1, 3, 4, and 6, respectively, based on the TMHMM algorithm of Sonnhammer, von Heijne, and Krogh (Erik L.L. Sonnhammer, Gunnar von Heijne, and Anders Krogh: A hidden Markov model for predicting transmembrane helices in protein sequences. In Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182 Ed J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen Menlo Park, CA: AAAI Press, 1998). The TMpred and TMHMM algorithms are accessed from the ExPasy molecular biology server located on the World Wide Web at (.expasy.ch/tools/). Protein variants 1 and 3 are predicted to contain 1 transmembrane region and protein variants 3 and 4 are not predicted to have transmembrane regions. All variants contain a hydrophobic stretch at their amino terminus that may encode a signal peptide.

Figure 17. Schematic alignment of SNP variants of 158P1D7. Schematic alignment of SNP variants of 158P1D7. Variant 158P1D7 v.2 is a variant with single nucleotide differences at 1546. Though this SNP variant is shown on transcript variant 158P1D7 v.1, it could also occur in any other transcript variants that contains the base pairs. Numbers correspond to those of 158P1D7 v.1. Black box shows sequence similar to 158P1D7 v.1. SNP is indicated above the box.

Figure 18. Schematic alignment of protein variants of 158P1D7. Schematic alignment of protein variants of 158P1D7. Protein variants correspond to nucleotide variants. Nucleotide variant 158P1D7 v.2 and v.5 code for the same protein as v.1. Nucleotide variants 158P1D7 v.3 and v.4 are transcript variants of v.1, as shown in Figure 12. Variant v.6 is a single nucleotide different from v.4 but codes for a protein that differs in the C-terminal portion from the protein coded by v.4. Black boxes represent sequence similar to v.1. Hatched box represents amino acid sequence not present in v.1. Numbers underneath the box correspond to 158P1D7 v.1.

Figure 19. Exon compositions of transcript variants of 158P1D7. Variant 158P1D7 v.3, v.4, v.5 and v.6 are transcript variants of 158P1D7 v.1. Variant 158P1D7 v.3 spliced 2069-2395 out of variant 158P1D7 v.1 and variant v.4 spliced out 1162-2096 out of v.1. Variant v.5 added another exon and 2 bp to the 5' end and extended 288 bp to the 3' end of variant v.1. Variant v.6 spliced at the same site as v.4 but spliced out an extra 'g' at the boundary. Numbers in "( )" underneath the boxes correspond to those of 158P1D7 v.1. Lengths of introns and exons are not proportional.

Figure 20. 158P1D7 Expression in Melanoma Cancer. RNA was extracted from normal skin cell line Detroit-551, and from the melanoma cancer cell line A375. Northern blots with 10ug of total RNA were probed with the 158P1D7 DNA probe. Size standards in kilobases are on the side. Results show expression of 158P1D7 in the melanoma cancer cell line but not in the normal cell line.

Figure 21. 158P1D7 Expression in cervical cancer patient specimens. First strand cDNA was prepared from normal cervix, cervical cancer cell line Hela, and a panel of cervical cancer patient specimens. Normalization was performed by PCR using primers to actin and GAPDH. Semi-quantitative PCR, using primers to 158P1D7, was performed at 26 and 30 cycles of amplification. Results show expression of 158P1D7 in 5 out of 14 tumor specimens tested but not in normal cervix nor in the cell line.

Figure 22. Detection of 158P1D7 protein in recombinant cells with monoclonal antibodies. Cell lysates from the indicated cell lines were separated by SDS-PAGE and then transferred to nitrocellulose for Western blotting. The blots were probed with 5 ug/ml of the indicated anti-158P1D7 monoclonal antibodies (MAbs) in PBS + 0.2% Tween 20 + 1% non-fat milk, washed, and then incubated with goat anti-mouse IgG-HRP secondary Ab. Immunoreactive bands were then

visualized by enhanced chemoluminescence and exposure to autoradiographic film. Arrows indicate the ~95 KD and 90 kD 158P1D7 protein doublet band which suggest 158P1D7 is post-translationally modified to generate 2 different molecular weight species. These results demonstrate expression of 158P1D7 protein in recombinant cells and specific detection of the protein with monoclonal antibodies.

Figure 23. Surface staining of 158P1D7-expressing 293T and UMUC cells with anti-158P1D7 monoclonal antibodies. Transiently transfected 293T cells expressing 158P1D7 and stable 158P1D7-expressing UMUC bladder cancer cells were analyzed for surface expression of 158P1D7 with monoclonal antibodies (MAbs) by flow cytometry. Transfected 293T control vector and 158P1D7 vector cells and stable UMUC-neo and UMUC-158P1D7 cells were stained with 10 ug/ml and 1 ug/ml, respectively, of the indicated MAbs. Surface bound MAbs were detected by incubation with goat anti-mouse IgG-PE secondary Ab and then subjected to FACS analysis. 158P1D7-expressing 293T and UMUC cells exhibited an increase in relative fluorescence comnpared to control cells demonstrating surface expression and detection of 158P1D7 protein by each of the MAbs.

Figure 24. Surface staining of endogenous 158P1D7-expressing LAPC9 prostate cancer and UGB1 bladder cancer xenograft cells with MAb M15-68(2)22.1.1. LAPC9 and UGB1 xenograft cells were subjected to surface staining with either control mouse IgG antibody or MAb M15-68(2).1.1 at 1 ug/ml. Surface bound MAbs were detected by incubation with goat anti-mouse IgG-PE secondary Ab and then subjected to FACS analysis. Both LAPC9 and UGB1 cells exhibited an increase in relative fluorescence with the anti-158P1D7 MAb demonstrating surface expression and detection of 158P1D7 protein.

Figure 25. Monoclonal antibody-mediated internalization of endogenous surface 158P1D7 in NCI-H146 small cell lung cancer cells. NCI-H146 cells were stained with 5 ug/ml of the indicated MAbs at 4°C for 1.5 hours, washed, and then either left at 4°C or moved to 37°C for 10 and 30 minutes. Residual surface bound MAb was then detected with anti-mouse IgG-PE secondary antibody. The decrease in the mean fluorescence intensity (MF) of cells moved to 37°C compared to cells left at 4°C demonstrates internalization of surface bound 158P1D7/MAb complexes.

Figrue 26. Binding of the 158P1D7 extracellular domain to human umbilical vein endothelial cells. The recombinant extracellular domain (ECD) of 158P1D7 (amino acids 16-608) was iodinated to high specific activity using the iodogen (1,3,4,5-tetrachloro-3a,6a-diphenylglycoluril) method. Human umbilical vein endothelial cells (HUVEC) at 90% confluency in 6 well plates was incubated with 1 nM of 125I-158P1D7 ECD in the presence (non-specific binding) or absence (Total binding) of 50 fold excess unlabeled ECD for 2 hours at either 4°C or 37°C. Cells were washed, solubilized in 0.5M NaOH, and subjected to gamma counting. The data shows specific binding of 158P1D7 ECD to HUVEC cells suggesting the presence of an 158P1D7 receptor on HUVEC cells. **Figure 26A.** Shows that the158P1D7 ECD bound directly to the surface of HUVEC cells as detected by the 158P1D7 specific MAb. **Figure 26B.** Shows specific binding of 158P1D7 ECD to HUVEC cells suggesting the presence of an 158P1D7 receptor on HUVEC cells.

Figure 27. 158P1D7 enhances the growth of bladder cancer in mice. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and maintained in a strictly controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals. 158P1D7 transfected UM-UC-3 cells and parental cells were injected into the subcutaneous space of SCID mice. Each mouse received 4 x 10^6 cells suspended in 50% (v/v) of Matrigel. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula $W^2 \times L/2$. The Mann-Whitney U test was used to evaluate differences of tumor growth. All tests were two sided with α=0.05.

Figure 28. Internalization of M15-68(2).31.1.1 in NCI-H146 cells. Endogenous-158P1D7 expressing NCI-H146 cells were incubated with 5 ug/ml of MAb M15-68(2).31.1.1 at 4°C for 1 hour, washed, and then incubated with goat anti-mouse IgG-PE secondary antibody and washed. Cells were then either left at 4°C or moved to 37°C for 30 minutes. Cells

were then subjected to fluorescent and brightfield microscopy. Cells that remained at 4°C exhibited a halo of fluorescence on the cells demonstrative of surface staining. Cells moved to 37°C exhibited a loss of the halo of surface fluorescence and the generation of punctate internal fluorescence indicative of internalization of the 158P1D7/MAb complexes.

Figure 29. Effect of 158P1D7 RNAi on cell survival. As control, 3T3 cells, a cell line with no detectable expression of 158P1D7 mRNA, was also treated with the panel of siRNAs (including oligo 158P1D7.b) and no phenotype was observed. This result reflects the fact that the specific protein knockdown in the LNCaP and PC3 cells is not a function of general toxicity, since the 3T3 cells did not respond to the 158P1D7.b oligo. The differential response of the three cell lines to the Eg5 control is a reflection of differences in levels of cell transfection and responsiveness of the cell lines to oligo treatment.

Figure 30. 158P1D7 MAb Retards the Growth of Human Bladder Cancer Xenografts in Mice. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and were maintained in a strictly-controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals. UG-B1, a patient bladder cancer, was used to establish xenograft models. Stock tumors regularly maintained in SCID mice were sterilely dissected, minced, and digested using Pronase (Calbiochem, San Diego, CA). Cell suspensions generated were incubated overnight at 37°C to obtain a homogeneous single-cell suspension. Each mouse received 2.5 x 10^6 cells at the subcutaneous site of right flank. A Murine monoclonal antibody to 158P1D7 was tested at a dose of 500 µg/mouse in the study. PBS was used as control. MAbs were dosed intra-peritoneally twice a week for a total of 12 doses, starting on the same day of tumor cell injection. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula: $W^2 \times L/2$. The results show that Anti-158P1D7 mAbs are capable of inhibiting the growth of human bladder carcinoma in mice.

Figure 31. 158P1D7 MAbs Retard Growth of Human Prostate Cancer Xenografts in Mice. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and were maintained in a strictly-controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals. LAPC-9AD, an androgen-dependent human prostate cancer, was used to establish xenograft models. Stock tumors were regularly maintained in SCID mice. At the day of implantation, stock tumors were harvested and trimmed of necrotic tissues and minced to 1 mm^3 pieces. Each mouse received 4 pieces of tissues at the subcutaneous site of right flank. A Murine monoclonal antibody to 158P1D7 was tested at a dose of 500 µg/mouse and 500 µg/mouse respectively. PBS and anti-KLH monoclonal antibody were used as controls. The study cohort consisted of 4 groups with 6 mice in each group. MAbs were dosed intra-peritoneally twice a week for a total of 8 doses. Treatment was started when tumor volume reached 45 mm^3. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula: $W^2 \times L/2$. The Student's t test and the Mann-Whitney U test, where applicable, were used to evaluate differences of tumor growth. All tests were two-sided with $\alpha=0.05$.

Figure 32. Effect of 158P1D7 on Proliferation of Rat1 cells. cells were grown overnight in 0.5% FBS and then compared to cells treated with 10% FBS. The cells were evaluated for proliferation at 18-96 hr post-treatment by a ^{3}H-thymidine incorporation assay and for cell cycle analysis by a BrdU incorporation/propidium iodide staining assay. The results show that the Rat-1 cells expressing the 158P1D7 antigen grew effectively in low serum concentrations (0.1%) compared to the Rat-1-Neo cells.

Figure 33. 158P1D7 Enhances Entry Into the S Phase. Cells were labeled with 10 µM BrdU, washed, trypsinized and fixed in 0.4% paraformaldehyde and 70% ethanol. Anti-BrdU-FITC (Pharmigen) was added to the cells, the cells were washed and then incubated with 10 µg/ml propidium iodide for 20 min prior to washing and analysis for fluorescence at 488 nm. The results show that there was increased labeling of cells in S-phase (DNA synthesis phase of the cell cycle) in 3T3 cells that expressed the 158P1D7 antigen relative to control cells.

Figure 34. Figure 34A. The cDNA (SEQ ID NO: 108) and amino acid sequence (SEQ ID NO: 109) of M15/X68(2)18 VH clone #1. **Figure 34B.** The cDNA (SEQ ID NO: 110) and amino acid sequence (SEQ ID NO: 111) of M15/X68(2)18 VL clone #2.

Figure 35. Figure 35A. The amino acid sequence (SEQ ID NO: 112) of M15/X68(2)18 VH clone #1. **Figure 35B.** The amino acid sequence (SEQ ID NO: 113) of M15/X68(2)18 VL clone #2.

Figure 36. Detection of 158P1D7 protein by immunohistochemistry in various cancer patient specimens. Tissue was obtained from patients with bladder transitional cell carcinoma, breast ductal carcinoma and lung carcinoma. The results showed expression of 158P1D7 in the tumor cells of the cancer patients' tissue panel (A) bladder transitional cell carcinoma, invasive Grade III (B) bladder transitional cell carcinoma, papillary Grade II. (C) breast infiltrating ductal carcinoma,moderately differentiated, (D)breast infiltrating ductal carcinoma,moderate to poorly differentiated, (E) lung squamous cell carcinoma, (F) lung adenocarcinoma, well differentiated. The expression of 158P1D7 in bladder transitional cell carcinoma tissues was detected mostly around the cell membrane indicating that 158P1D7 is membrane associated.

DETAILED DESCRIPTION OF THE INVENTION

Outline of Sections

I.) Definitions:

Unless otherwise defined, all terms of art, notations and other scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this invention pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. Many of the techniques and procedures described or referenced herein are well understood and commonly employed using conventional methodology by those skilled in the art, such as, for example, the widely utilized molecular cloning methodologies described in Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd. edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. As appropriate, procedures involving the use of commercially available kits and reagents are generally carried out in accordance with manufacturer defined protocols and/or parameters unless otherwise noted.

The terms "invasive bladder cancer" means bladder cancers that have extended into the bladder muscle wall, and are meant to include stage stage T2 - T4 and disease under the TNM (tumor, node, metastasis) system. In general, these patients have substantially less favorable outcomes compared to patients having non-invasive cancer. Following cystectomy, 50% or more of the patients with invasive cancer will develop metastasis (Whittmore. Semin Urol 1983; 1:4-10).

"Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence 158P1D7 (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence 158P1D7. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

The term "analog" refers to a molecule which is structurally similar or shares similar or corresponding attributes with another molecule (e.g. a 158P1D7-related protein). For example an analog of the 158P1D7 protein can be specifically bound by an antibody or T cell that specifically binds to 158P1D7 protein.

The term "antibody" is used in the broadest sense. Therefore an "antibody" can be naturally occurring or man-made such as monoclonal antibodies produced by conventional hybridoma technology. Anti-158P1D7 antibodies bind 158P1D7 proteins, or a fragment thereof, and comprise monoclonal and polyclonal antibodies as well as fragments containing the antigen-binding domain and/or one or more complementarity determining regions of these antibodies.

An "antibody fragment" is defined as at least a portion of the variable region of the immunoglobulin molecule that binds to its target, i.e., the antigen-binding region. In one embodiment it specifically covers single anti-158P1D7 antibodies and

clones thereof (including agonist, antagonist and neutralizing antibodies) and anti-158P1D7 antibody compositions with polyepitopic specificity.

The term "codon optimized sequences" refers to nucleotide sequences that have been optimized for a particular host species by replacing any one or more than one codon having a usage frequency of less than about 20%, more preferably less than about 30% or 40%. A sequence may be "completely optimized" to contain no codon having a usage frequency of less than about 20%, more preferably less than about 30% or 40%. Nucleotide sequences that have been optimized for expression in a given host species by elimination of spurious polyadenylation sequences, elimination of exon/intron splicing signals, elimination of transposon-like repeats and/or optimization of GC content in addition to codon optimization are referred to herein as an "expression enhanced sequences."

The term "cytotoxic agent" refers to a substance that inhibits or prevents one or more than one function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes chemotherapeutic agents, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof. Examples of cytotoxic agents include, but are not limited to maytansinoids, yttrium, bismuth, ricin, ricin A-chain, doxorubicin, daunorubicin, taxol, ethidium bromide, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, dihydroxy anthracin dione, actinomycin, diphtheria toxin, Pseudomonas exotoxin (PE) A, PE40, abrin, abrin A chain, modeccin A chain, alpha-sarcin, gelonin, mitogellin, retstrictocin, phenomycin, enomycin, curicin, crotin, calicheamicin, sapaonaria officinalis inhibitor, and glucocorticoid and other chemotherapeutic agents, as well as radioisotopes such as At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32} and radioactive isotopes of Lu. Antibodies may also be conjugated to an anti-cancer pro-drug activating enzyme capable of converting the pro-drug to its active form.

The term "homolog" refers to a molecule which exhibits homology to another molecule, by for example, having sequences of chemical residues that are the same or similar at corresponding positions.

"Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see, e.g.*, Stites, *et al.*, IMMUNOLOGY, 8TH ED., Lange Publishing, Los Altos, CA (1994).

The terms "hybridize", "hybridizing", "hybridizes" and the like, used in the context of polynucleotides, are meant to refer to conventional hybridization conditions, preferably such as hybridization in 50% formamide/6XSSC/0.1% SDS/100 µg/ml ssDNA, in which temperatures for hybridization are above 37 degrees C and temperatures for washing in 0.1XSSC/0.1% SDS are above 55 degrees C.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated, or present, with the peptides in their *in situ* environment. For example, a polynucleotide is said to be "isolated" when it is substantially separated from contaminant polynucleotides that correspond or are complementary to nucleic acids other than those of 158P1D7 or that encode polypeptides other than 158P1D7 gene product or fragments thereof. A skilled artisan can readily employ nucleic acid isolation procedures to obtain an isolated 158P1D7 polynucleotide. A protein is said to be "isolated," for example, when physical, mechanical and/or chemical methods are employed to remove the 158P1D7 protein from cellular constituents that are normally associated, or present, with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated 158P1D7 protein. Alternatively, an isolated protein can be prepared by synthetic or chemical means.

The term "mammal" refers to any organism classified as a mammal, including mice, rats, rabbits, dogs, cats, cows, horses and humans. In one embodiment of the invention, the mammal is a mouse. In another embodiment of the invention, the mammal is a human.

The terms "metastatic bladder cancer" and "metastatic disease" mean bladder cancers that have spread to regional lymph nodes or to distant sites, and are meant to stage TxNxM+ under the TNM system. The most common site for bladder cancer metastasis is lymph node. Other common sites for metastasis include lung, bone and liver.

The term "monoclonal antibody" refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the antibodies comprising the population are identical except for possible naturally occurring mutations that are present in minor amounts.

A "motif", as in biological motif of an 158P1D7-related protein, refers to any pattern of amino acids forming part of the primary sequence of a protein, that is associated with a particular function (e.g. protein-protein interaction, protein-DNA interaction, etc) or modification (e.g. that is phosphorylated, glycosylated or amidated), or localization (e.g. secretory sequence, nuclear localization sequence, etc.) or a sequence that is correlated with being immunogenic, either humorally or cellularly. A motif can be either contiguous or capable of being aligned to certain positions that are generally correlated with a certain function or property. In the context of HLA motifs, "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs for HLA binding are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "pharmaceutical excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservative, and the like.

"Pharmaceutically acceptable" refers to a non-toxic, inert, and/or composition that is physiologically compatible with mammals, such as humans.

The term "polynucleotide" means a polymeric form of nucleotides of at least 3, 4, 5, 6, 7, 8, 9, or 10 bases or base pairs in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide, and is meant to include single and double stranded forms of DNA and/or RNA. In the art, this term is often used interchangeably with "oligonucleotide", although "oligonucleotide" may be used to refer to the subset of polynucleotides less than about 50 nucleotides in length. A polynucleotide can comprise a nucleotide sequence disclosed herein wherein thymidine (T) (as shown for example in can also be uracil (U); this definition pertains to the differences between the chemical structures of DNA and RNA, in particular the observation that one of the four major bases in RNA is uracil (U) instead of thymidine (T).

The term "polypeptide" means a polymer of at least about 4, 5, 6, 7, or 8 amino acids. Throughout the specification, standard three letter or single letter designations for amino acids are used. In the art, this term is often used interchangeably with "peptide" or "protein", thus "peptide" may be used to refer to the subset of polypeptides less than about 50 amino acids in length.

An HLA "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding groove of an HLA molecule, with their side chains buried in specific pockets of the binding groove. In one embodiment, for example, the primary anchor residues for an HLA class I molecule are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 8, 9, 10, 11, or 12 residue peptide epitope in accordance with the invention. In another embodiment, for example, the primary anchor residues of a peptide that will bind an HLA class II molecule are spaced relative to each other, rather than to the termini of a peptide, where the peptide is generally of at least 9 amino acids in length. The primary anchor positions for each motif and supermotif are set forth in Table IV. For example, analog peptides can be created by altering the presence or absence of

particular residues in the primary and/or secondary anchor positions shown in Table IV. Such analogs are used to modulate the binding affinity and/or population coverage of a peptide comprising a particular HLA motif or supermotif.

A "recombinant" DNA or RNA molecule is a DNA or RNA molecule that has been subjected to molecular manipulation *in vitro*.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured nucleic acid sequences to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, are identified by, but not limited to, those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42 °C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42 °C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium. citrate) and 50% formamide at 55 °C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55 °C. "Moderately stringent conditions" are described by, but not limited to, those in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

An HLA "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles.

A "transgenic animal" (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A "transgene" is a DNA that is integrated into the genome of a cell from which a transgenic animal develops.

As used herein, an HLA or cellular immune response "vaccine" is a composition that contains or encodes one or more peptides of the invention. There are numerous embodiments of such vaccines, such as a cocktail of one or more individual peptides; one or more peptides of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such individual peptides or polypeptides, *e.g.*, a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include any whole unit integer from 1-150 or more, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 or more peptides of the invention. The peptides or polypeptides can optionally be modified, such as by lipidation, addition of targeting or other sequences. HLA class I peptides of the invention can be admixed with, or linked to, HLA class II peptides, to facilitate activation of both

cytotoxic T lymphocytes and helper T lymphocytes. HLA vaccines can also comprise peptide-pulsed antigen presenting cells, *e.g.*, dendritic cells.

The term "variant" refers to a molecule that exhibits a variation from a described type or norm, such as a protein that has one or more different amino acid residues in the corresponding position(s) of a specifically described protein (e.g. the 158P1D7 protein shown in Figure 2 or Figure 3). An analog is an example of a variant protein.

The 158P1D7-related proteins of the invention include those specifically identified herein, as well as allelic variants, conservative substitution variants, analogs and homologs that can be isolated/generated and characterized without undue experimentation following the methods outlined herein or readily available in the art. Fusion proteins that combine parts of different 158P1D7 proteins or fragments thereof, as well as fusion proteins of a 158P1D7 protein and a heterologous polypeptide are also included. Such 158P1D7 proteins are collectively referred to as the 158P1D7-related proteins, the proteins of the invention, or 158P1D7. The term "158P1D7-related protein" refers to a polypeptide fragment or an 158P1D7 protein sequence of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or more than 25 amino acids; or, at least about 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100 or more than 100 amino acids.

II.) 158P1D7 Polynucleotides

One aspect of the invention provides polynucleotides corresponding or complementary to all or part of an 158P1D7 gene, mRNA, and/or coding sequence, preferably in isolated form, including polynucleotides encoding an 158P1D7-related protein and fragments thereof, DNA, RNA, DNA/RNA hybrid, and related molecules, polynucleotides or oligonucleotides complementary to an 158P1D7 gene or mRNA sequence or a part thereof, and polynucleotides or oligonucleotides that hybridize to an 158P1D7 gene, mRNA, or to an 158P1D7 encoding polynucleotide (collectively, "158P1D7 polynucleotides"). In all instances when referred to in this section, T can also be U in Figure 2.

Embodiments of a 158P1D7 polynucleotide include: a 158P1D7 polynucleotide having the sequence shown in Figure 2, the nucleotide sequence of 158P1D7 as shown in Figure 2, wherein T is U; at least 10 contiguous nucleotides of a polynucleotide having the sequence as shown in Figure 2; or, at least 10 contiguous nucleotides of a polynucleotide having the sequence as shown in Figure 2 where T is U. For example, embodiments of 158P1D7 nucleotides comprise, without limitation:

(a) a polynucleotide comprising or consisting of the sequence as shown in Figure 2, wherein T can also be U;

(b) a polynucleotide comprising or consisting of the sequence as shown in Figure 2, from nucleotide residue number 23 through nucleotide residue number 2548, wherein T can also be U;

(c) a polynucleotide that encodes a 158P1D7-related protein whose sequence is encoded by the cDNAs contained in the plasmid designated p158P1D7- Turbo/3PX deposited with American Type Culture Collection as Accession No. PTA-3662 on 22 August 2001 (sent via Federal Express on 20 August 2001);

(d) a polynucleotide that encodes an 158P1D7-related protein that is at least 90% homologous to the entire amino acid sequence shown in Figure 2;

(e) a polynucleotide that encodes an 158P1D7-related protein that is at least 90% identical to the entire amino acid sequence shown in Figure 2;

(f) a polynucleotide that encodes at least one peptide set forth in Tables V-XVIII;

(g) a polynucleotide that encodes a peptide region of at least 5 amino acids of Figure 3 in any whole number increment up to 841 that includes an amino acid position having a value greater than 0.5 in the Hydrophilicity profile of Figure 11;

(h) a polynucleotide that encodes a peptide region of at least 5 amino acids of Figure 3 in any whole number increment up to 841 that includes an amino acid position having a value less than 0.5 in the Hydropathicity profile of Figure 12;

(i) a polynucleotide that encodes a peptide region of at least 5 amino acids of Figure 3 in any whole number increment up to 841 that includes an amino acid position having a value greater than 0.5 in the Percent Accessible Residues profile of Figure 13;

(j) a polynucleotide that encodes a peptide region of at least 5 amino acids of Figure 3 in any whole number increment up to 841 that includes an amino acid position having a value greater than 0.5 in the Average Flexibility profile on Figure 14;

(k) a polynucleotide that encodes a peptide region of at least 5 amino acids of Figure 3 in any whole number increment up to 841 that includes an amino acid position having a value greater than 0.5 in the Beta-turn profile of Figure 15;

(l) a polynucleotide that is fully complementary to a polynucleotide of any one of (a)-(k);

(m) a polynucleotide that selectively hybridizes under stringent conditions to a polynucleotide of (a)-(l);

(n) a peptide that is encoded by any of (a)–(k); and,

(o) a polynucleotide of any of (a)-(m) or peptide of (n) together with a pharmaceutical excipient and/or in a human unit dose form.

As used herein, a range is understood to specifically disclose all whole unit positions thereof.

Typical embodiments of the invention disclosed herein include 158P1D7 polynucleotides that encode specific portions of the 158P1D7 mRNA sequence (and those which are complementary to such sequences) such as those that encode the protein and fragments thereof, for example of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825 or 841 contiguous amino acids.

For example, representative embodiments of the invention disclosed herein include: polynucleotides and their encoded peptides themselves encoding about amino acid 1 to about amino acid 10 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 10 to about amino acid 20 of the 158P1D7 protein shown in Figure 2, or Figure 3, polynucleotides encoding about amino acid 20 to about amino acid 30 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 30 to about amino acid 40 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 40 to about amino acid 50 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 50 to about amino acid 60 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 60 to about amino acid 70 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 70 to about amino acid 80 of the 158P1D7 protein shown in

Figure 2 or Figure 3, polynucleotides encoding about amino acid 80 to about amino acid 90 of the 158P1D7 protein shown in Figure 2 or Figure 3, polynucleotides encoding about amino acid 90 to about amino acid 100 of the 158P1D7 protein shown in Figure 2 or Figure 3, in increments of about 10 amino acids, ending at the carboxyl terminal amino acid set forth in Figure 2 or Figure 3. Accordingly polynucleotides encoding portions of the amino acid sequence (of about 10 amino acids), of amino acids 100 through the carboxyl terminal amino acid of the 158P1D7 protein are embodiments of the invention. Wherein it is understood that each particular amino acid position discloses that position plus or minus five amino acid residues.

Polynucleotides encoding relatively long portions of the 158P1D7 protein are also within the scope of the invention. For example, polynucleotides encoding from about amino acid 1 (or 20 or 30 or 40 etc.) to about amino acid 20, (or 30, or 40 or 50 etc.) of the 158P1D7 protein shown in Figure 2 or Figure 3 can be generated by a variety of techniques well known in the art. These polynucleotide fragments can include any portion of the 158P1D7 sequence as shown in Figure 2 or Figure 3.

Additional illustrative embodiments of the invention disclosed herein include 158P1D7 polynucleotide fragments encoding one or more of the biological motifs contained within the 158P1D7 protein sequence, including one or more of the motif-bearing subsequences of the 158P1D7 protein set forth in Tables V-XVIII. In another embodiment, typical polynucleotide fragments of the invention encode one or more of the regions of 158P1D7 that exhibit homology to a known molecule. In another embodiment of the invention, typical polynucleotide fragments can encode one or more of the 158P1D7 N-glycosylation sites, cAMP and cGMP-dependent protein kinase phosphorylation sites, casein kinase II phosphorylation sites or N-myristoylation site and amidation sites.

II.A.) Uses of 158P1D7 Polynucleotides

II.A.1.) Monitoring of Genetic Abnormalities

The polynucleotides of the preceding paragraphs have a number of different specific uses. The human 158P1D7 gene maps to the chromosomal location set forth in Example 3. For example, because the 158P1D7 gene maps to this chromosome, polynucleotides that encode different regions of the 158P1D7 protein are used to characterize cytogenetic abnormalities of this chromosomal locale, such as abnormalities that are identified as being associated with various cancers. In certain genes, a variety of chromosomal abnormalities including rearrangements have been identified as frequent cytogenetic abnormalities in a number of different cancers (see e.g. Krajinovic et al., Mutat. Res. 382(3-4): 81-83 (1998); Johansson et al., Blood 86(10): 3905-3914 (1995) and Finger et al., P.N.A.S. 85(23): 9158-9162 (1988)). Thus, polynucleotides encoding specific regions of the 158P1D7 protein provide new tools that can be used to delineate, with greater precision than previously possible, cytogenetic abnormalities in the chromosomal region that encodes 158P1D7 that may contribute to the malignant phenotype. In this context, these polynucleotides satisfy a need in the art for expanding the sensitivity of chromosomal screening in order to identify more subtle and less common chromosomal abnormalities (see e.g. Evans et al., Am. J. Obstet. Gynecol 171(4): 1055-1057 (1994)).

Furthermore, as 158P1D7 was shown to be highly expressed in bladder and other cancers, 158P1D7 polynucleotides are used in methods assessing the status of 158P1D7 gene products in normal versus cancerous tissues. Typically, polynucleotides that encode specific regions of the 158P1D7 protein are used to assess the presence of perturbations (such as deletions, insertions, point mutations, or alterations resulting in a loss of an antigen etc.) in specific regions of the 158P1D7 gene, such as such regions containing one or more motifs. Exemplary assays include both RT-PCR assays as well as single-strand conformation polymorphism (SSCP) analysis (see, e.g., Marrogi et al., J. Cutan. Pathol. 26(8): 369-378 (1999), both of which utilize polynucleotides encoding specific regions of a protein to examine these regions within the protein.

II.A.2.) Antisense Embodiments

Other specifically contemplated nucleic acid related embodiments of the invention disclosed herein are genomic DNA, cDNAs, ribozymes, and antisense molecules, as well as nucleic acid molecules based on an alternative backbone, or including alternative bases, whether derived from natural sources or synthesized, and include molecules capable of inhibiting the RNA or protein expression of 158P1D7. For example, antisense molecules can be RNAs or other molecules, including peptide nucleic acids (PNAs) or non-nucleic acid molecules such as phosphorothioate derivatives, that specifically bind DNA or RNA in a base pair-dependent manner. A skilled artisan can readily obtain these classes of nucleic acid molecules using the 158P1D7 polynucleotides and polynucleotide sequences disclosed herein.

Antisense technology entails the administration of exogenous oligonucleotides that bind to a target polynucleotide located within the cells. The term "antisense" refers to the fact that such oligonucleotides are complementary to their intracellular targets, e.g., 158P1D7. See for example, Jack Cohen, Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression, CRC Press, 1989; and Synthesis 1:1-5 (1988). The 158P1D7 antisense oligonucleotides of the present invention include derivatives such as S-oligonucleotides (phosphorothioate derivatives or S-oligos, see, Jack Cohen, supra), which exhibit enhanced cancer cell growth inhibitory action. S-oligos (nucleoside phosphorothioates) are isoelectronic analogs of an oligonucleotide (O-oligo) in which a nonbridging oxygen atom of the phosphate group is replaced by a sulfur atom. The S-oligos of the present invention can be prepared by treatment of the corresponding O-oligos with 3H-1,2-benzodithiol-3-one-1,1-dioxide, which is a sulfur transfer reagent. See Iyer, R. P. et al, J. Org. Chem. 55:4693-4698 (1990); and Iyer, R. P. et al., J. Am. Chem. Soc. 112:1253-1254 (1990). Additional 158P1D7 antisense oligonucleotides of the present invention include morpholino antisense oligonucleotides known in the art (see, e.g., Partridge et al., 1996, Antisense & Nucleic Acid Drug Development 6: 169-175).

The 158P1D7 antisense oligonucleotides of the present invention typically can be RNA or DNA that is complementary to and stably hybridizes with the first 100 5' codons or last 100 3' codons of the 158P1D7 genomic sequence or the corresponding mRNA. Absolute complementarity is not required, although high degrees of complementarity are preferred. Use of an oligonucleotide complementary to this region allows for the selective hybridization to 158P1D7 mRNA and not to mRNA specifying other regulatory subunits of protein kinase. In one embodiment, 158P1D7 antisense oligonucleotides of the present invention are 15 to 30-mer fragments of the antisense DNA molecule that have a sequence that hybridizes to 158P1D7 mRNA. Optionally, 158P1D7 antisense oligonucleotide is a 30-mer oligonucleotide that is complementary to a region in the first 10 5' codons or last 10 3' codons of 158P1D7. Alternatively, the antisense molecules are modified to employ ribozymes in the inhibition of 158P1D7 expression, see, e.g., L. A. Couture & D. T. Stinchcomb; *Trends Genet* 12: 510-515 (1996).

II.A.3.) Primers and Primer Pairs

Further specific embodiments of this nucleotides of the invention include primers and primer pairs, which allow the specific amplification of polynucleotides of the invention or of any specific parts thereof, and probes that selectively or specifically hybridize to nucleic acid molecules of the invention or to any part thereof. Primers may also be used as probes and can be labeled with a detectable marker, such as, for example, a radioisotope, fluorescent compound, bioluminescent compound, a chemiluminescent compound, metal chelator or enzyme. Such probes and primers are used to detect the presence of a 158P1D7 polynucleotide in a sample and as a means for detecting a cell expressing a 158P1D7 protein.

Examples of such probes include polypeptides comprising all or part of the human 158P1D7 cDNA sequence shown in Figure 2. Examples of primer pairs capable of specifically amplifying 158P1D7 mRNAs are also described in the Examples. As will be understood by the skilled artisan, a great many different primers and probes can be prepared based on the sequences provided herein and used effectively to amplify and/or detect a 158P1D7 mRNA. Preferred probes of the invention are polynucleotides of more than about 9, about 12, about 15, about 18, about 20, about 23, about 25, about 30, about 35, about 40, about 45, and about 50 consecutive nucleotides found in 158P1D7 nucleic acids disclosed herein.

The 158P1D7 polynucleotides of the invention are useful for a variety of purposes, including but not limited to their use as probes and primers for the amplification and/or detection of the 158P1D7 gene(s), mRNA(s), or fragments thereof; as reagents for the diagnosis and/or prognosis of bladder cancer and other cancers; as coding sequences capable of directing the expression of 158P1D7 polypeptides; as tools for modulating or inhibiting the expression of the 158P1D7 gene(s) and/or translation of the 158P1D7 transcript(s); and as therapeutic agents.

II.A.4.) Isolation of 158P1D7-Encoding Nucleic Acid Molecules

The 158P1D7 cDNA sequences described herein enable the isolation of other polynucleotides encoding 158P1D7 gene product(s), as well as the isolation of polynucleotides encoding 158P1D7 gene product homologs, alternatively spliced isoforms, allelic variants, and mutant forms of the 158P1D7 gene product as well as polynucleotides that encode analogs of 158P1D7-related proteins. Various molecular cloning methods that can be employed to isolate full length cDNAs encoding an 158P1D7 gene are well known (see, for example, Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2d edition, Cold Spring Harbor Press, New York; 1989; Current Protocols in Molecular Biology. Ausubel et al., Eds., Wiley and Sons, 1995). For example, lambda phage cloning methodologies can be conveniently employed, using commercially available cloning systems (e.g., Lambda ZAP Express, Stratagene). Phage clones containing 158P1D7 gene cDNAs can be identified by probing with a labeled 158P1D7 cDNA or a fragment thereof. For example, in one embodiment, the 158P1D7 cDNA (Figure 2) or a portion thereof can be synthesized and used as a probe to retrieve overlapping and full-length cDNAs corresponding to a 158P1D7 gene. The 158P1D7 gene itself can be isolated by screening genomic DNA libraries, bacterial artificial chromosome libraries (BACs), yeast artificial chromosome libraries (YACs), and the like, with 158P1D7 DNA probes or primers.

The present invention includes the use of any probe as described herein to identify and isolate a 158P1D7 or 158P1D7 related nucleic acid sequence from a naturally occurring source, such as humans or other mammals, as well as the isolated nucleic acid sequence *per se*, which would comprise all or most of the sequences found in the probe used.

II.A.5.) Recombinant Nucleic Acid Molecules and Host-Vector Systems

The invention also provides recombinant DNA or RNA molecules containing an 158P1D7 polynucleotide, a fragment, analog or homologue thereof, including but not limited to phages, plasmids, phagemids, cosmids, YACs, BACs, as well as various viral and non-viral vectors well known in the art, and cells transformed or transfected with such recombinant DNA or RNA molecules. Methods for generating such molecules are well known (see, for example, Sambrook et al, 1989, supra).The invention further provides a host-vector system comprising a recombinant DNA molecule containing a 158P1D7 polynucleotide, fragment, analog or homologue thereof within a suitable prokaryotic or eukaryotic host cell. Examples of suitable eukaryotic host cells include a yeast cell, a plant cell, or an animal cell, such as a mammalian cell or an insect cell (e.g., a baculovirus-infectible cell such as an Sf9 or HighFive cell). Examples of suitable mammalian cells include various bladder cancer cell lines such as SCaBER, UM-UC3, HT1376, RT4, T24, TCC-SUP, J82 and SW780, other transfectable or transducible bladder cancer cell lines, as well as a number of mammalian cells routinely used for the expression of recombinant proteins (e.g., COS, CHO, 293, 293T cells). More particularly, a polynucleotide comprising the coding sequence of 158P1D7 or a fragment, analog or homolog thereof can be used to generate 158P1D7 proteins or fragments thereof using any number of host-vector systems routinely used and widely known in the art.

A wide range of host-vector systems suitable for the expression of 158P1D7 proteins or fragments thereof are available, see for example, Sambrook et al., 1989, supra; Current Protocols in Molecular Biology, 1995, supra). Preferred vectors for mammalian expression include but are not limited to pcDNA 3.1 myc-His-tag (Invitrogen) and the retroviral vector pSRαtkneo (Muller et al., 1991, MCB 11:1785). Using these expression vectors, 158P1D7 can be expressed in several bladder cancer and non-bladder cell lines, including for example SCaBER, UM-UC3, HT1376, RT4, T24, TCC-SUP, J82 and SW780. The host-vector systems of the invention are useful for the production of a 158P1D7 protein or fragment thereof.

Such host-vector systems can be employed to study the functional properties of 158P1D7 and 158P1D7 mutations or analogs.

Recombinant human 158P1D7 protein or an analog or homolog or fragment thereof can be produced by mammalian cells transfected with a construct encoding a 158P1D7-related nucleotide. For example, 293T cells can be transfected with an expression plasmid encoding 158P1D7 or fragment, analog or homolog thereof, the 158P1D7 or related protein is expressed in the 293T cells, and the recombinant 158P1D7 protein is isolated using standard purification methods (e.g., affinity purification using anti-158P1D7 antibodies). In another embodiment, a 158P1D7 coding sequence is subcloned into the retroviral vector pSRαMSVtkneo and used to infect various mammalian cell lines, such as NIH 3T3, TsuPr1, 293 and rat-1 in order to establish 158P1D7 expressing cell lines. Various other expression systems well known in the art can also be employed. Expression constructs encoding a leader peptide joined in frame to the 158P1D7 coding sequence can be used for the generation of a secreted form of recombinant 158P1D7 protein.

As discussed herein, redundancy in the genetic code permits variation in 158P1D7 gene sequences. In particular, it is known in the art that specific host species often have specific codon preferences, and thus one can adapt the disclosed sequence as preferred for a desired host. For example, preferred analog codon sequences typically have rare codons (i.e., codons having a usage frequency of less than about 20% in known sequences of the desired host) replaced with higher frequency codons. Codon preferences for a specific species are calculated, for example, by utilizing codon usage tables available on the INTERNET such as at URL URL: dna.affrc.go.jp/~nakamura/codon.html.

Additional sequence modifications are known to enhance protein expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon/intron splice site signals, transposon-like repeats, and/or other such well-characterized sequences that are deleterious to gene expression. The GC content of the sequence is adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. Where possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures. Other useful modifications include the addition of a translational initiation consensus sequence at the start of the open reading frame, as described in Kozak, *Mol. Cell Biol.*, 9:5073-5080 (1989). Skilled artisans understand that the general rule that eukaryotic ribosomes initiate translation exclusively at the 5' proximal AUG codon is abrogated only under rare conditions (see, e.g., Kozak PNAS 92(7): 2662-2666, (1995) and Kozak NAR 15(20): 8125-8148 (1987)).

III.) 158P1D7-related Proteins

Another aspect of the present invention provides 158P1D7-related proteins. Specific embodiments of 158P1D7 proteins comprise a polypeptide having all or part of the amino acid sequence of human 158P1D7 as shown in Figure 2 or Figure 3. Alternatively, embodiments of 158P1D7 proteins comprise variant, homolog or analog polypeptides that have alterations in the amino acid sequence of 158P1D7 shown in Figure 2 or Figure 3.

In general, naturally occurring allelic variants of human 158P1D7 share a high degree of structural identity and homology (e.g., 90% or more homology). Typically, allelic variants of the 158P1D7 protein contain conservative amino acid substitutions within the 158P1D7 sequences described herein or contain a substitution of an amino acid from a corresponding position in a homologue of 158P1D7. One class of 158P1D7 allelic variants are proteins that share a high degree of homology with at least a small region of a particular 158P1D7 amino acid sequence, but further contain a radical departure from the sequence, such as a non-conservative substitution, truncation, insertion or frame shift. In comparisons of protein sequences, the terms, similarity, identity, and homology each have a distinct meaning as appreciated in the field of genetics. Moreover, orthology and paralogy can be important concepts describing the relationship of members of a given protein family in one organism to the members of the same family in other organisms.

Amino acid abbreviations are provided in Table II. Conservative amino acid substitutions can frequently be made in a protein without altering either the conformation or the function of the protein. Proteins of the invention can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more conservative substitutions. Such changes include substituting any of isoleucine (I), valine (V), and leucine (L) for any other of these hydrophobic amino acids; aspartic acid (D) for glutamic acid (E) and vice versa; glutamine (Q) for asparagine (N) and vice versa; and serine (S) for threonine (T) and vice versa. Other substitutions can also be considered conservative, depending on the environment of the particular amino acid and its role in the three-dimensional structure of the protein. For example, glycine (G) and alanine (A) can frequently be interchangeable, as can alanine (A) and valine (V). Methionine (M), which is relatively hydrophobic, can frequently be interchanged with leucine and isoleucine, and sometimes with valine. Lysine (K) and arginine (R) are frequently interchangeable in locations in which the significant feature of the amino acid residue is its charge and the differing pK's of these two amino acid residues are not significant. Still other changes can be considered "conservative" in particular environments (see, e.g. Table III herein; pages 13-15 "Biochemistry" 2nd ED. Lubert Stryer ed (Stanford University); Henikoff et al., PNAS 1992 Vol 89 10915-10919; Lei et al., J Biol Chem 1995 May 19; 270(20):11882-6).

Embodiments of the invention disclosed herein include a wide variety of art-accepted variants or analogs of 158P1D7 proteins such as polypeptides having amino acid insertions, deletions and substitutions. 158P1D7 variants can be made using methods known in the art such as site-directed mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis (Carter et al., *Nucl. Acids Res., 13*:4331 (1986); Zoller et al., *Nucl. Acids Res., 10*:6487 (1987)), cassette mutagenesis (Wells et al., Gene, 34:315 (1985)), restriction selection mutagenesis (Wells et al., *Philos. Trans. R. Soc. London SerA*, 317:415 (1986)) or other known techniques can be performed on the cloned DNA to produce the 158P1D7 variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence that is involved in a specific biological activity such as a protein-protein interaction. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions (Creighton, *The Proteins*, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)). If alanine substitution does not yield adequate amounts of variant, an isosteric amino acid can be used.

As defined herein, 158P1D7 variants, analogs or homologs, have the distinguishing attribute of having at least one epitope that is "cross reactive" with a 158P1D7 protein having the amino acid sequence of Figure 2. As used in this sentence, "cross reactive" means that an antibody or T cell that specifically binds to an 158P1D7 variant also specifically binds to the 158P1D7 protein having the amino acid sequence of Figure 2. A polypeptide ceases to be a variant of the protein shown in Figure 2 when it no longer contains any epitope capable of being recognized by an antibody or T cell that specifically binds to the 158P1D7 protein. Those skilled in the art understand that antibodies that recognize proteins bind to epitopes of varying size, and a grouping of the order of about four or five amino acids, contiguous or not, is regarded as a typical number of amino acids in a minimal epitope. See, e.g., Nair et al., J. Immunol 2000 165(12): 6949-6955; Hebbes et al., Mol Immunol (1989) 26(9):865-73; Schwartz et al., J Immunol (1985) 135(4):2598-608.

Another class of 158P1D7-related protein variants share 70%, 75%, 80%, 85% or 90% or more similarity with the amino acid sequence of Figure 2 or a fragment thereof. Another specific class of 158P1D7 protein variants or analogs comprise one or more of the 158P1D7 biological motifs described herein or presently known in the art. Thus, encompassed by the present invention are analogs of 158P1D7 fragments (nucleic or amino acid) that have altered functional (e.g.

immunogenic) properties relative to the starting fragment. It is to be appreciated that motifs now or which become part of the art are to be applied to the nucleic or amino acid sequences of Figure 2 or Figure 3.

As discussed herein, embodiments of the claimed invention include polypeptides containing less than the full amino acid sequence of the 158P1D7 protein shown in Figure 2 or Figure 3. For example, representative embodiments of the invention comprise peptides/proteins having any 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more contiguous amino acids of the 158P1D7 protein shown in Figure 2 or Figure 3.

Moreover, representative embodiments of the invention disclosed herein include polypeptides consisting of about amino acid 1 to about amino acid 10 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 10 to about amino acid 20 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 20 to about amino acid 30 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 30 to about amino acid 40 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 40 to about amino acid 50 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 50 to about amino acid 60 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 60 to about amino acid 70 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 70 to about amino acid 80 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 80 to about amino acid 90 of the 158P1D7 protein shown in Figure 2 or Figure 3, polypeptides consisting of about amino acid 90 to about amino acid 100 of the 158P1D7 protein shown in Figure 2 or Figure 3, etc. throughout the entirety of the 158P1D7 amino acid sequence. Moreover, polypeptides consisting of about amino acid 1 (or 20 or 30 or 40 etc.) to about amino acid 20, (or 130, or 140 or 150 etc.) of the 158P1D7 protein shown in Figure 2 or Figure 3 are embodiments of the invention. It is to be appreciated that the starting and stopping positions in this paragraph refer to the specified position as well as that position plus or minus 5 residues.

158P1D7-related proteins are generated using standard peptide synthesis technology or using chemical cleavage methods well known in the art. Alternatively, recombinant methods can be used to generate nucleic acid molecules that encode a 158P1D7-related protein. In one embodiment, nucleic acid molecules provide a means to generate defined fragments of the 158P1D7 protein (or variants, homologs or analogs thereof).

III.A.) Motif-bearing Protein Embodiments

Additional illustrative embodiments of the invention disclosed herein include 158P1D7 polypeptides comprising the amino acid residues of one or more of the biological motifs contained within the 158P1D7 polypeptide sequence set forth in Figure 2 or Figure 3. Various motifs are known in the art, and a protein can be evaluated for the presence of such motifs by a number of publicly available Internet sites (see, e.g., URL addresses: pfam.wustl.edu/; searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html; psort.ims.u-tokyo.ac.jp/; URL: cbs.dtu.dk/; ebi.ac.uk/interpro/scan.html; expasy.ch/tools/scnpsit1.html; Epimatrix™ and Epimer™, Brown University, brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html; and BIMAS, bimas.dcrt.nih.gov/.).

Motif bearing subsequences of the 158P1D7 protein are set forth and identified in Table XIX.

Table XX sets forth several frequently occurring motifs based on pfam searches (see URL address pfam.wustl.edu/). The columns of Table XX list (1) motif name abbreviation, (2) percent identity found amongst the different member of the motif family, (3) motif name or description and (4) most common function; location information is included if the motif is relevant for location.

Polypeptides comprising one or more of the 158P1D7 motifs discussed above are useful in elucidating the specific characteristics of a malignant phenotype in view of the observation that the 158P1D7 motifs discussed above are associated with growth dysregulation and because 158P1D7 is overexpressed in certain cancers (See, e.g., Table I). Casein kinase II, cAMP and camp-dependent protein kinase, and Protein Kinase C, for example, are enzymes known to be associated with

the development of the malignant phenotype (see e.g. Chen et al., Lab Invest., 78(2): 165-174 (1998); Gaiddon et al., Endocrinology 136(10): 4331-4338 (1995); Hall et al., Nucleic Acids Research 24(6): 1119-1126 (1996); Peterziel et al., Oncogene 18(46): 6322-6329 (1999) and O'Brian, Oncol. Rep. 5(2): 305-309 (1998)). Moreover, both glycosylation and myristoylation are protein modifications also associated with cancer and cancer progression (see e.g. Dennis et al., Biochem. Biophys. Acta 1473(1):21-34 (1999); Raju et al., Exp. Cell Res. 235(1): 145-154 (1997)). Amidation is another protein modification also associated with cancer and cancer progression (see e.g. Treston et al., J. Natl. Cancer Inst. Monogr. (13): 169-175 (1992)).

In another embodiment, proteins of the invention comprise one or more of the immunoreactive epitopes identified in accordance with art-accepted methods, such as the peptides set forth in Tables V-XVIII. CTL epitopes can be determined using specific algorithms to identify peptides within an 158P1D7 protein that are capable of optimally binding to specified HLA alleles (e.g., Table IV; Epimatrix™ and Epimer™, Brown University, URL: brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html; and BIMAS, URL: bimas.dcrt.nih.gov/.) Moreover, processes for identifying peptides that have sufficient binding affinity for HLA molecules and which are correlated with being immunogenic epitopes, are well known in the art, and are carried out without undue experimentation. In addition, processes for identifying peptides that are immunogenic epitopes, are well known in the art, and are carried out without undue experimentation either *in vitro* or *in vivo*.

Also known in the art are principles for creating analogs of such epitopes in order to modulate immunogenicity. For example, one begins with an epitope that bears a CTL or HTL motif (see, e.g., the HLA Class I and HLA Class II motifs/supermotifs of Table IV). The epitope is analoged by substituting out an amino acid at one of the specified positions, and replacing it with another amino acid specified for that position. For example, one can substitute out a deleterious residue in favor of any other residue, such as a preferred residue as defined in Table IV; substitute a less-preferred residue with a preferred residue as defined in Table IV; or substitute an originally-occurring preferred residue with another preferred residue as defined in Table IV. Substitutions can occur at primary anchor positions or at other positions in a peptide; see, e.g., Table IV.

A variety of references reflect the art regarding the identification and generation of epitopes in a protein of interest as well as analogs thereof. See, for example, WO 9733602 to Chesnut et al.; Sette, Immunogenetics 1999 50(3-4): 201-212; Sette et al., J. Immunol. 2001 166(2): 1389-1397; Sidney et al., Hum. Immunol. 1997 58(1): 12-20; Kondo et al., Immunogenetics 1997 45(4): 249-258; Sidney et al., J. Immunol. 1996 157(8): 3480-90; and Falk et al., Nature 351: 290-6 (1991); Hunt et al., Science 255:1261-3 (1992); Parker et al., J. Immunol. 149:3580-7 (1992); Parker et al., J. Immunol. 152:163-75 (1994)); Kast et al., 1994 152(8): 3904-12; Borras-Cuesta et al., Hum. Immunol. 2000 61(3): 266-278; Alexander et al., J. Immunol. 2000 164(3); 164(3): 1625-1633; Alexander et al., PMID: 7895164, UI: 95202582; O'Sullivan et al., J. Immunol. 1991 147(8): 2663-2669; Alexander et al., Immunity 1994 1(9): 751-761 and Alexander et al., Immunol. Res. 1998 18(2): 79-92.

Related embodiments of the inventions include polypeptides comprising combinations of the different motifs set forth in Table XIX, and/or, one or more of the predicted CTL epitopes of Table V through Table XVIII, and/or, one or more of the T cell binding motifs known in the art. Preferred embodiments contain no insertions, deletions or substitutions either within the motifs or the intervening sequences of the polypeptides. In addition, embodiments which include a number of either N-terminal and/or C-terminal amino acid residues on either side of these motifs may be desirable (to, for example, include a greater portion of the polypeptide architecture in which the motif is located). Typically the number of N-terminal and/or C-terminal amino acid residues on either side of a motif is between about 1 to about 100 amino acid residues, preferably 5 to about 50 amino acid residues.

158P1D7-related proteins are embodied in many forms, preferably in isolated form. A purified 158P1D7 protein molecule will be substantially free of other proteins or molecules that impair the binding of 158P1D7 to antibody, T cell or other ligand. The nature and degree of isolation and purification will depend on the intended use. Embodiments of a 158P1D7-related proteins include purified 158P1D7-related proteins and functional, soluble 158P1D7-related proteins. In one embodiment, a functional, soluble 158P1D7 protein or fragment thereof retains the ability to be bound by antibody, T cell or other ligand.

The invention also provides 158P1D7 proteins comprising biologically active fragments of the 158P1D7 amino acid sequence shown in Figure 2 or Figure 3. Such proteins exhibit properties of the 158P1D7 protein, such as the ability to elicit the generation of antibodies that specifically bind an epitope associated with the 158P1D7 protein; to be bound by such antibodies; to elicit the activation of HTL or CTL; and/or, to be recognized by HTL or CTL.

158P1D7-related polypeptides that contain particularly interesting structures can be predicted and/or identified using various analytical techniques well known in the art, including, for example, the methods of Chou-Fasman, Garnier-Robson, Kyte-Doolittle, Eisenberg, Karplus-Schultz or Jameson-Wolf analysis, or on the basis of immunogenicity. Fragments that contain such structures are particularly useful in generating subunit-specific anti-158P1D7 antibodies, or T cells or in identifying cellular factors that bind to 158P1D7.

CTL epitopes can be determined using specific algorithms to identify peptides within an 158P1D7 protein that are capable of optimally binding to specified HLA alleles (e.g., by using the SYFPEITHI site at World Wide Web URL syfpeithi.bmi-heidelberg.com/; the listings in Table IV(A)-(E); Epimatrix™ and Epimer™, Brown University, URL (URL: brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html); and BIMAS, URL: bimas.dcrt.nih.gov/). Illustrating this, peptide epitopes from 158P1D7 that are presented in the context of human MHC class I molecules HLA-A1, A2, A3, A11, A24, B7 and B35 were predicted (Tables V-XVIII). Specifically, the complete amino acid sequence of the 158P1D7 protein was entered into the HLA Peptide Motif Search algorithm found in the Bioinformatics and Molecular Analysis Section (BIMAS) web site listed above. The HLA peptide motif search algorithm was developed by Dr. Ken Parker based on binding of specific peptide sequences in the groove of HLA Class I molecules, in particular HLA-A2 (see, e.g., Falk et al., Nature 351: 290-6 (1991); Hunt et al., Science 255:1261-3 (1992); Parker et al., J. Immunol. 149:3580-7 (1992); Parker et al., J. Immunol. 152:163-75 (1994)). This algorithm allows location and ranking of 8-mer, 9-mer, and 10-mer peptides from a complete protein sequence for predicted binding to HLA-A2 as well as numerous other HLA Class I molecules. Many HLA class I binding peptides are 8-, 9-, 10 or 11-mers. For example, for class I HLA-A2, the epitopes preferably contain a leucine (L) or methionine (M) at position 2 and a valine (V) or leucine (L) at the C-terminus (see, e.g., Parker et al., J. Immunol. 149:3580-7 (1992)). Selected results of 158P1D7 predicted binding peptides are shown in Tables V-XVIII herein. In Tables V-XVIII, the top 50 ranking candidates, 9-mers and 10-mers, for each family member are shown along with their location, the amino acid sequence of each specific peptide, and an estimated binding score. The binding score corresponds to the estimated half time of dissociation of complexes containing the peptide at 37°C at pH 6.5. Peptides with the highest binding score are predicted to be the most tightly bound to HLA Class I on the cell surface for the greatest period of time and thus represent the best immunogenic targets for T-cell recognition.

Actual binding of peptides to an HLA allele can be evaluated by stabilization of HLA expression on the antigen-processing defective cell line T2 (see, e.g., Xue et al., Prostate 30:73-8 (1997) and Peshwa et al., Prostate 36:129-38 (1998)). Immunogenicity of specific peptides can be evaluated *in vitro* by stimulation of CD8+ cytotoxic T lymphocytes (CTL) in the presence of antigen presenting cells such as dendritic cells.

It is to be appreciated that every epitope predicted by the BIMAS site, Epimer™ and Epimatrix™ sites, or specified by the HLA class I or class II motifs available in the art or which become part of the art such as set forth in Table IV (or

determined using World Wide Web site URL syfpeithi.bmi-heidelberg.com/) are to be "applied" to the 158P1D7 protein. As used in this context "applied" means that the 158P1D7 protein is evaluated, e.g., visually or by computer-based patterns finding methods, as appreciated by those of skill in the relevant art. Every subsequence of the 158P1D7 of 8, 9, 10, or 11 amino acid residues that bears an HLA Class I motif, or a subsequence of 9 or more amino acid residues that bear an HLA Class II motif are within the scope of the invention.

III.B.) Expression of 158P1D7-related Proteins

In an embodiment described in the examples that follow, 158P1D7 can be conveniently expressed in cells (such as 293T cells) transfected with a commercially available expression vector such as a CMV-driven expression vector encoding 158P1D7 with a C-terminal 6XHis and MYC tag (pcDNA3.1/mycHIS, Invitrogen or Tag5, GenHunter Corporation, Nashville TN). The Tag5 vector provides an IgGK secretion signal that can be used to facilitate the production of a secreted 158P1D7 protein in transfected cells. The secreted HIS-tagged 158P1D7 in the culture media can be purified, e.g., using a nickel column using standard techniques.

III.C.) Modifications of 158P1D7-related Proteins

Modifications of 158P1D7-related proteins such as covalent modifications are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a 158P1D7 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the 158P1D7. Another type of covalent modification of the 158P1D7 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of a protein of the invention. Another type of covalent modification of 158P1D7 comprises linking the 158P1D7 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The 158P1D7-related proteins of the present invention can also be modified to form a chimeric molecule comprising 158P1D7 fused to another, heterologous polypeptide or amino acid sequence. Such a chimeric molecule can be synthesized chemically or recombinantly. A chimeric molecule can have a protein of the invention fused to another tumor-associated antigen or fragment thereof. Alternatively, a protein in accordance with the invention can comprise a fusion of fragments of the 158P1D7 sequence (amino or nucleic acid) such that a molecule is created that is not, through its length, directly homologous to the amino or nucleic acid sequences shown in Figure 2 or Figure 3. Such a chimeric molecule can comprise multiples of the same subsequence of 158P1D7. A chimeric molecule can comprise a fusion of a 158P1D7-related protein with a polyhistidine epitope tag, which provides an epitope to which immobilized nickel can selectively bind, with cytokines or with growth factors. The epitope tag is generally placed at the amino- or carboxyl- terminus of the 158P1D7. In an alternative embodiment, the chimeric molecule can comprise a fusion of a 158P1D7-related protein with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a 158P1D7 polypeptide in place of at least one variable region within an Ig molecule. In a preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CHI, CH2 and CH3 regions of an IgGI molecule. For the production of immunoglobulin fusions see, e.g., U.S. Patent No. 5,428,130 issued June 27, 1995.

III.D.) Uses of 158P1D7-related Proteins

The proteins of the invention have a number of different uses. As 158P1D7 is highly expressed in bladder and other cancers, 158P1D7-related proteins are used in methods that assess the status of 158P1D7 gene products in normal versus cancerous tissues, thereby elucidating the malignant phenotype. Typically, polypeptides from specific regions of the 158P1D7 protein are used to assess the presence of perturbations (such as deletions, insertions, point mutations etc.) in

those regions (such as regions containing one or more motifs). Exemplary assays utilize antibodies or T cells targeting 158P1D7-related proteins comprising the amino acid residues of one or more of the biological motifs contained within the 158P1D7 polypeptide sequence in order to evaluate the characteristics of this region in normal versus cancerous tissues or to elicit an immune response to the epitope. Alternatively, 158P1D7-related proteins that contain the amino acid residues of one or more of the biological motifs in the 158P1D7 protein are used to screen for factors that interact with that region of 158P1D7.

158P1D7 protein fragments/subsequences are particularly useful in generating and characterizing domain-specific antibodies (e.g., antibodies recognizing an extracellular or intracellular epitope of an 158P1D7 protein), for identifying agents or cellular factors that bind to 158P1D7 or a particular structural domain thereof, and in various therapeutic and diagnostic contexts, including but not limited to diagnostic assays, cancer vaccines and methods of preparing such vaccines.

Proteins encoded by the 158P1D7 genes, or by analogs, homologs or fragments thereof, have a variety of uses, including but not limited to generating antibodies and in methods for identifying ligands and other agents and cellular constituents that bind to an 158P1D7 gene product. Antibodies raised against an 158P1D7 protein or fragment thereof are useful in diagnostic and prognostic assays, and imaging methodologies in the management of human cancers characterized by expression of 158P1D7 protein, such as those listed in Table I. Such antibodies can be expressed intracellularly and used in methods of treating patients with such cancers. 158P1D7-related nucleic acids or proteins are also used in generating HTL or CTL responses.

Various immunological assays useful for the detection of 158P1D7 proteins are used, including but not limited to various types of radioimmunoassays, enzyme-linked immunosorbent assays (ELISA), enzyme-linked immunofluorescent assays (ELIFA), immunocytochemical methods, and the like. Antibodies can be labeled and used as immunological imaging reagents capable of detecting 158P1D7-expressing cells (e.g., in radioscintigraphic imaging methods). 158P1D7 proteins are also particularly useful in generating cancer vaccines, as further described herein.

IV.) 158P1D7 Antibodies

Another aspect of the invention provides antibodies that bind to 158P1D7-related proteins. Preferred antibodies specifically bind to a 158P1D7-related protein and do not bind (or bind weakly) to peptides or proteins that are not 158P1D7-related proteins. For example, antibodies bind 158P1D7 can bind 158P1D7-related proteins such as the homologs or analogs thereof.

158P1D7 antibodies of the invention are particularly useful in bladder cancer diagnostic and prognostic assays, and imaging methodologies. Similarly, such antibodies are useful in the treatment, diagnosis, and/or prognosis of other cancers, to the extent 158P1D7 is also expressed or overexpressed in these other cancers. Moreover, intracellularly expressed antibodies (e.g., single chain antibodies) are therapeutically useful in treating cancers in which the expression of 158P1D7 is involved, such as advanced or metastatic bladder cancers.

The invention also provides various immunological assays useful for the detection and quantification of 158P1D7 and mutant 158P1D7-related proteins. Such assays can comprise one or more 158P1D7 antibodies capable of recognizing and binding a 158P1D7-related protein, as appropriate. These assays are performed within various immunological assay formats well known in the art, including but not limited to various types of radioimmunoassays, enzyme-linked immunosorbent assays (ELISA), enzyme-linked immunofluorescent assays (ELIFA), and the like.

Immunological non-antibody assays of the invention also comprise T cell immunogenicity assays (inhibitory or stimulatory) as well as major histocompatibility complex (MHC) binding assays.

In addition, immunological imaging methods capable of detecting bladder cancer and other cancers expressing 158P1D7 are also provided by the invention, including but not limited to radioscintigraphic imaging methods using labeled

158P1D7 antibodies. Such assays are clinically useful in the detection, monitoring, and prognosis of 158P1D7 expressing cancers such as bladder cancer.

158P1D7 antibodies are also used in methods for purifying a 158P1D7-related protein and for isolating 158P1D7 homologues and related molecules. For example, a method of purifying a 158P1D7-related protein comprises incubating an 158P1D7 antibody, which has been coupled to a solid matrix, with a lysate or other solution containing a 158P1D7-related protein under conditions that permit the 158P1D7 antibody to bind to the 158P1D7-related protein; washing the solid matrix to eliminate impurities; and eluting the 158P1D7-related protein from the coupled antibody. Other uses of the 158P1D7 antibodies of the invention include generating anti-idiotypic antibodies that mimic the 158P1D7 protein.

Various methods for the preparation of antibodies are well known in the art. For example, antibodies can be prepared by immunizing a suitable mammalian host using a 158P1D7-related protein, peptide, or fragment, in isolated or immunoconjugated form (Antibodies: A Laboratory Manual, CSH Press, Eds., Harlow, and Lane (1988); Harlow, Antibodies, Cold Spring Harbor Press, NY (1989)). In addition, fusion proteins of 158P1D7 can also be used, such as a 158P1D7 GST-fusion protein. In a particular embodiment, a GST fusion protein comprising all or most of the amino acid sequence of Figure 2 or Figure 3 is produced, then used as an immunogen to generate appropriate antibodies. In another embodiment, a 158P1D7-related protein is synthesized and used as an immunogen.

In addition, naked DNA immunization techniques known in the art are used (with or without purified 158P1D7-related protein or 158P1D7 expressing cells) to generate an immune response to the encoded immunogen (for review, see Donnelly et al., 1997, Ann. Rev. Immunol. 15: 617-648).

The amino acid sequence of 158P1D7 as shown in Figure 2 or Figure 3 can be analyzed to select specific regions of the 158P1D7 protein for generating antibodies. For example, hydrophobicity and hydrophilicity analyses of the 158P1D7 amino acid sequence are used to identify hydrophilic regions in the 158P1D7 structure (see, e. g., the Example entitled "Antigenicity profiles"). Regions of the 158P1D7 protein that show immunogenic structure, as well as other regions and domains, can readily be identified using various other methods known in the art, such as Chou-Fasman, Hopp and Woods, Kyte-Doolittle, Janin, Bhaskaran and Ponnuswamy, Deleage and Roux, Garnier-Robson, Eisenberg, Karplus-Schultz, or Jameson-Wolf analysis. Thus, each region identified by any of these programs or methods is within the scope of the present invention. Methods for the generation of 158P1D7 antibodies are further illustrated by way of the examples provided herein. Methods for preparing a protein or polypeptide for use as an immunogen are well known in the art. Also well known in the art are methods for preparing immunogenic conjugates of a protein with a carrier, such as BSA, KLH or other carrier protein. In some circumstances, direct conjugation using, for example, carbodiimide reagents are used; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, IL, are effective. Administration of a 158P1D7 immunogen is often conducted by injection over a suitable time period and with use of a suitable adjuvant, as is understood in the art. During the immunization schedule, titers of antibodies can be taken to determine adequacy of antibody formation.

158P1D7 monoclonal antibodies can be produced by various means well known in the art. For example, immortalized cell lines that secrete a desired monoclonal antibody are prepared using the standard hybridoma technology of Kohler and Milstein or modifications that immortalize antibody-producing B cells, as is generally known. Immortalized cell lines that secrete the desired antibodies are screened by immunoassay in which the antigen is a 158P1D7-related protein. When the appropriate immortalized cell culture is identified, the cells can be expanded and antibodies produced either from *in vitro* cultures or from ascites fluid.

One embodiment of the invention is a mouse hybridoma that produces murine monoclonal antibodies designated X68(2)18 (a.k.a. M15-68(2)18.1.1) deposited with American Type Culture Collection (ATCC), P.O. Box 1549, Manassas, VA 20108 on 06-February-2004 and assigned Accession No. _____.

The antibodies or fragments of the invention can also be produced, by recombinant means. Regions that bind specifically to the desired regions of the 158P1D7 protein can also be produced in the context of chimeric or complementarity determining region (CDR) grafted antibodies of multiple species origin. Humanized or human 158P1D7 antibodies can also be produced, and are preferred for use in therapeutic contexts. Methods for humanizing murine and other non-human antibodies, by substituting one or more of the non-human antibody CDRs for corresponding human antibody sequences, are well known (see for example, Jones et al., 1986, Nature 321: 522-525; Riechmann et al., 1988, Nature 332: 323-327; Verhoeyen et al., 1988, Science 239: 1534-1536). See also, Carter et al., 1993, Proc. Natl. Acad. Sci. USA 89: 4285 and Sims et al., 1993, J. Immunol. 151: 2296.

Methods for producing fully human monoclonal antibodies include phage display and transgenic methods (for review, see Vaughan et al., 1998, Nature Biotechnology 16: 535-539). Fully human 158P1D7 monoclonal antibodies can be generated using cloning technologies employing large human Ig gene combinatorial libraries (i.e., phage display) (Griffiths and Hoogenboom, Building an *in vitro* immune system: human antibodies from phage display libraries. In: Protein Engineering of Antibody Molecules for Prophylactic and Therapeutic Applications in Man, Clark, M. (Ed.), Nottingham Academic, pp 45-64 (1993); Burton and Barbas, Human Antibodies from combinatorial libraries. Id., pp 65-82). Fully human 158P1D7 monoclonal antibodies can also be produced using transgenic mice engineered to contain human immunoglobulin gene loci as described in PCT Patent Application WO98/24893, Kucherlapati and Jakobovits et al., published December 3, 1997 (see also, Jakobovits, 1998, Exp. Opin. Invest. Drugs 7(4): 607-614; U.S. patents 6,162,963 issued 19 December 2000; 6,150,584 issued 12 November 2000; and, 6,114598 issued 5 September 2000). This method avoids the *in vitro* manipulation required with phage display technology and efficiently produces high affinity authentic human antibodies.

Reactivity of 158P1D7 antibodies with an 158P1D7-related protein can be established by a number of well known means, including Western blot, immunoprecipitation, ELISA, and FACS analyses using, as appropriate, 158P1D7-related proteins, 158P1D7-expressing cells or extracts thereof. A 158P1D7 antibody or fragment thereof can be labeled with a detectable marker or conjugated to a second molecule. Suitable detectable markers include, but are not limited to, a radioisotope, a fluorescent compound, a bioluminescent compound, chemiluminescent compound, a metal chelator or an enzyme. Further, bi-specific antibodies specific for two or more 158P1D7 epitopes are generated using methods generally known in the art. Homodimeric antibodies can also be generated by cross-linking techniques known in the art (e.g., Wolff et al., Cancer Res. 53: 2560-2565).

V.) 158P1D7 Cellular Immune Responses

The mechanism by which T cells recognize antigens has been delineated. Efficacious peptide epitope vaccine compositions of the invention induce a therapeutic or prophylactic immune responses in very broad segments of the world-wide population. For an understanding of the value and efficacy of compositions of the invention that induce cellular immune responses, a brief review of immunology-related technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. *et al.*, *Cell* 47:1071, 1986; Babbitt, B. P. *et al.*, *Nature* 317:359, 1985; Townsend, A. and Bodmer, H., *Annu. Rev. Immunol.* 7:601, 1989; Germain, R. N., *Annu. Rev. Immunol.* 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are set forth in Table IV (*see* also, *e.g.*, Southwood, *et al.*, *J. Immunol.* 160:3363, 1998; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995; Rammensee *et al.*, SYFPEITHI, access via World Wide Web at URL syfpeithi.bmi-heidelberg.com/; Sette, A. and Sidney, J. *Curr. Opin. Immunol.* 10:478, 1998; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994; Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992; Sinigaglia, F. and Hammer, J. *Curr. Biol.* 6:52, 1994; Ruppert *et al.*, *Cell* 74:929-937, 1993; Kondo *et*

al., J. Immunol. 155:4307-4312, 1995; Sidney *et al., J. Immunol.* 157:3480-3490, 1996; Sidney *et al., Human Immunol.* 45:79-93, 1996; Sette, A. and Sidney, J. *Immunogenetics 1999 Nov; 50(3-4):201-12,* Review).

Furthermore, x-ray crystallographic analyses of HLA-peptide complexes have revealed pockets within the peptide binding cleft/groove of HLA molecules which accommodate, in an allele-specific mode, residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present. (*See, e.g.,* Madden, D.R. *Annu. Rev. Immunol.* 13:587, 1995; Smith, *et al., Immunity* 4:203, 1996; Fremont *et al., Immunity* 8:305, 1998; Stern *et al., Structure* 2:245, 1994; Jones, E.Y. *Curr. Opin. Immunol.* 9:75, 1997; Brown, J. H. *et al., Nature* 364:33, 1993; Guo, H. C. *et al., Proc. Natl. Acad. Sci. USA* 90:8053, 1993; Guo, H. C. *et al., Nature* 360:364, 1992; Silver, M. L. *et al., Nature* 360:367, 1992; Matsumura, M. *et al., Science* 257:927, 1992; Madden *et al., Cell* 70:1035, 1992; Fremont, D. H. *et al., Science* 257:919, 1992; Saper, M. A. , Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991.)

Accordingly, the definition of class I and class II allele-specific HLA binding motifs, or class I or class II supermotifs allows identification of regions within a protein that are correlated with binding to particular HLA antigen(s).

Thus, by a process of HLA motif identification, candidates for epitope-based vaccines have been identified; such candidates can be further evaluated by HLA-peptide binding assays to determine binding affinity and/or the time period of association of the epitope and its corresponding HLA molecule. Additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of population coverage, and/or immunogenicity.

Various strategies can be utilized to evaluate cellular immunogenicity, including:

1) Evaluation of primary T cell cultures from normal individuals (*see, e.g.,* Wentworth, P. A. *et al., Mol. Immunol.* 32:603, 1995; Celis, E. *et al., Proc. Natl. Acad. Sci. USA* 91:2105, 1994; Tsai, V. *et al., J. Immunol.* 158:1796, 1997; Kawashima, I. *et al., Human Immunol.* 59:1, 1998). This procedure involves the stimulation of peripheral blood lymphocytes (PBL) from normal subjects with a test peptide in the presence of antigen presenting cells *in vitro* over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using, *e.g.,* a lymphokine- or ^{51}Cr-release assay involving peptide sensitized target cells.

2) Immunization of HLA transgenic mice (*see, e.g.,* Wentworth, P. A. *et al., J. Immunol.* 26:97, 1996; Wentworth, P. A. *et al., Int. Immunol.* 8:651, 1996; Alexander, J. *et al., J. Immunol.* 159:4753, 1997). For example, in such methods peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured *in vitro* in the presence of test peptide for approximately one week. Peptide-specific T cells are detected using, *e.g.,* a ^{51}Cr-release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.

3) Demonstration of recall T cell responses from immune individuals who have been either effectively vaccinated and/or from chronically ill patients (*see, e.g.,* Rehermann, B. *et al., J. Exp. Med.* 181:1047, 1995; Doolan, D. L. *et al., Immunity* 7:97, 1997; Bertoni, R. *et al., J. Clin. Invest.* 100:503, 1997; Threlkeld, S. C. *et al., J. Immunol.* 159:1648, 1997; Diepolder, H. M. *et al., J. Virol.* 71:6011, 1997). Accordingly, recall responses are detected by culturing PBL from subjects that have been exposed to the antigen due to disease and thus have generated an immune response "naturally", or from patients who were vaccinated against the antigen. PBL from subjects are cultured *in vitro* for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays including ^{51}Cr release involving peptide-sensitized targets, T cell proliferation, or lymphokine release.

VI.) 158P1D7 Transgenic Animals

Nucleic acids that encode a 158P1D7-related protein can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. In accordance with established techniques, cDNA encoding 158P1D7 can be used to clone genomic DNA that encodes 158P1D7. The cloned genomic sequences can then be used to generate transgenic animals containing cells that express DNA that encode 158P1D7. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 issued 12 April 1988, and 4,870,009 issued 26 September 1989. Typically, particular cells would be targeted for 158P1D7 transgene incorporation with tissue-specific enhancers.

Transgenic animals that include a copy of a transgene encoding 158P1D7 can be used to examine the effect of increased expression of DNA that encodes 158P1D7. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this aspect of the invention, an animal is treated with a reagent and a reduced incidence of a pathological condition, compared to untreated animals that bear the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of 158P1D7 can be used to construct a 158P1D7 "knock out" animal that has a defective or altered gene encoding 158P1D7 as a result of homologous recombination between the endogenous gene encoding 158P1D7 and altered genomic DNA encoding 158P1D7 introduced into an embryonic cell of the animal. For example, cDNA that encodes 158P1D7 can be used to clone genomic DNA encoding 158P1D7 in accordance with established techniques. A portion of the genomic DNA encoding 158P1D7 can be deleted or replaced with another gene, such as a gene encoding a selectable marker that can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector (see, e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected (see, e.g.,, Li et al., Cell, 69:915 (1992)). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras (see, e.g.,, Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal, and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knock out animals can be characterized, for example, for their ability to defend against certain pathological conditions or for their development of pathological conditions due to absence of the 158P1D7 polypeptide.

VII.) Methods for the Detection of 158P1D7

Another aspect of the present invention relates to methods for detecting 158P1D7 polynucleotides and polypeptides and 158P1D7-related proteins, as well as methods for identifying a cell that expresses 158P1D7. The expression profile of 158P1D7 makes it a diagnostic marker for metastasized disease. Accordingly, the status of 158P1D7 gene products provides information useful for predicting a variety of factors including susceptibility to advanced stage disease, rate of progression, and/or tumor aggressiveness. As discussed in detail herein, the status of 158P1D7 gene products in patient samples can be analyzed by a variety protocols that are well known in the art including immunohistochemical analysis, the variety of Northern blotting techniques including *in situ* hybridization, RT-PCR analysis (for example on laser capture micro-dissected samples), Western blot analysis and tissue array analysis.

More particularly, the invention provides assays for the detection of 158P1D7 polynucleotides in a biological sample, such as urine, serum, bone, prostatic fluid, tissues, semen, cell preparations, and the like. Detectable 158P1D7 polynucleotides

include, for example, a 158P1D7 gene or fragment thereof, 158P1D7 mRNA, alternative splice variant 158P1D7 mRNAs, and recombinant DNA or RNA molecules that contain a 158P1D7 polynucleotide. A number of methods for amplifying and/or detecting the presence of 158P1D7 polynucleotides are well known in the art and can be employed in the practice of this aspect of the invention.

In one embodiment, a method for detecting an 158P1D7 mRNA in a biological sample comprises producing cDNA from the sample by reverse transcription using at least one primer; amplifying the cDNA so produced using an 158P1D7 polynucleotides as sense and antisense primers to amplify 158P1D7 cDNAs therein; and detecting the presence of the amplified 158P1D7 cDNA. Optionally, the sequence of the amplified 158P1D7 cDNA can be determined.

In another embodiment, a method of detecting a 158P1D7 gene in a biological sample comprises first isolating genomic DNA from the sample; amplifying the isolated genomic DNA using 158P1D7 polynucleotides as sense and antisense primers; and detecting the presence of the amplified 158P1D7 gene. Any number of appropriate sense and antisense probe combinations can be designed from the nucleotide sequence provided for the 158P1D7 (Figure 2) and used for this purpose.

The invention also provides assays for detecting the presence of an 158P1D7 protein in a tissue or other biological sample such as urine, serum, semen, bone, prostate, cell preparations, and the like. Methods for detecting a 158P1D7-related protein are also well known and include, for example, immunoprecipitation, immunohistochemical analysis, Western blot analysis, molecular binding assays, ELISA, ELIFA and the like. For example, a method of detecting the presence of a 158P1D7-related protein in a biological sample comprises first contacting the sample with a 158P1D7 antibody, a 158P1D7-reactive fragment thereof, or a recombinant protein containing an antigen binding region of a 158P1D7 antibody; and then detecting the binding of 158P1D7-related protein in the sample.

Methods for identifying a cell that expresses 158P1D7 are also within the scope of the invention. In one embodiment, an assay for identifying a cell that expresses a 158P1D7 gene comprises detecting the presence of 158P1D7 mRNA in the cell. Methods for the detection of particular mRNAs in cells are well known and include, for example, hybridization assays using complementary DNA probes (such as *in situ* hybridization using labeled 158P1D7 riboprobes, Northern blot and related techniques) and various nucleic acid amplification assays (such as RT-PCR using complementary primers specific for 158P1D7, and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like). Alternatively, an assay for identifying a cell that expresses a 158P1D7 gene comprises detecting the presence of 158P1D7-related protein in the cell or secreted by the cell. Various methods for the detection of proteins are well known in the art and are employed for the detection of 158P1D7-related proteins and cells that express 158P1D7-related proteins.

158P1D7 expression analysis is also useful as a tool for identifying and evaluating agents that modulate 158P1D7 gene expression. For example, 158P1D7 expression is significantly upregulated in bladder cancer, and is expressed in cancers of the tissues listed in Table I. Identification of a molecule or biological agent that inhibits 158P1D7 expression or over-expression in cancer cells is of therapeutic value. For example, such an agent can be identified by using a screen that quantifies 158P1D7 expression by RT-PCR, nucleic acid hybridization or antibody binding.

VIII.) Methods for Monitoring the Status of 158P1D7-related Genes and Their Products

Oncogenesis is known to be a multistep process where cellular growth becomes progressively dysregulated and cells progress from a normal physiological state to precancerous and then cancerous states (see, e.g., Alers et al., Lab Invest. 77(5): 437-438 (1997) and Isaacs et al., Cancer Surv. 23: 19-32 (1995)). In this context, examining a biological sample for evidence of dysregulated cell growth (such as aberrant 158P1D7 expression in cancers) allows for early detection of such aberrant physiology, before a pathologic state such as cancer has progressed to a stage that therapeutic options are more limited and or the prognosis is worse. In such examinations, the status of 158P1D7 in a biological sample of interest

can be compared, for example, to the status of 158P1D7 in a corresponding normal sample (e.g. a sample from that individual or alternatively another individual that is not affected by a pathology). An alteration in the status of 158P1D7 in the biological sample (as compared to the normal sample) provides evidence of dysregulated cellular growth. In addition to using a biological sample that is not affected by a pathology as a normal sample, one can also use a predetermined normative value such as a predetermined normal level of mRNA expression (see, e.g., Grever et al., J. Comp. Neurol. 1996 Dec 9;376(2):306-14 and U.S. Patent No. 5,837,501) to compare 158P1D7 status in a sample.

The term "status" in this context is used according to its art accepted meaning and refers to the condition or state of a gene and its products. Typically, skilled artisans use a number of parameters to evaluate the condition or state of a gene and its products. These include, but are not limited to the location of expressed gene products (including the location of 158P1D7 expressing cells) as well as the level, and biological activity of expressed gene products (such as 158P1D7 mRNA, polynucleotides and polypeptides). Typically, an alteration in the status of 158P1D7 comprises a change in the location of 158P1D7 and/or 158P1D7 expressing cells and/or an increase in 158P1D7 mRNA and/or protein expression.

158P1D7 status in a sample can be analyzed by a number of means well known in the art, including without limitation, immunohistochemical analysis, *in situ* hybridization, RT-PCR analysis on laser capture micro-dissected samples, Western blot analysis, and tissue array analysis. Typical protocols for evaluating the status of the 158P1D7 gene and gene products are found, for example in Ausubel et al. eds., 1995, Current Protocols In Molecular Biology, Units 2 (Northern Blotting), 4 (Southern Blotting), 15 (Immunoblotting) and 18 (PCR Analysis). Thus, the status of 158P1D7 in a biological sample is evaluated by various methods utilized by skilled artisans including, but not limited to genomic Southern analysis (to examine, for example perturbations in the 158P1D7 gene), Northern analysis and/or PCR analysis of 158P1D7 mRNA (to examine, for example alterations in the polynucleotide sequences or expression levels of 158P1D7 mRNAs), and, Western and/or immunohistochemical analysis (to examine, for example alterations in polypeptide sequences, alterations in polypeptide localization within a sample, alterations in expression levels of 158P1D7 proteins and/or associations of 158P1D7 proteins with polypeptide binding partners). Detectable 158P1D7 polynucleotides include, for example, a 158P1D7 gene or fragment thereof, 158P1D7 mRNA, alternative splice variants, 158P1D7 mRNAs, and recombinant DNA or RNA molecules containing a 158P1D7 polynucleotide.

The expression profile of 158P1D7 makes it a diagnostic marker for local and/or metastasized disease, and provides information on the growth or oncogenic potential of a biological sample. In particular, the status of 158P1D7 provides information useful for predicting susceptibility to particular disease stages, progression, and/or tumor aggressiveness. The invention provides methods and assays for determining 158P1D7 status and diagnosing cancers that express 158P1D7, such as cancers of the tissues listed in Table I. For example, because 158P1D7 mRNA is so highly expressed in bladder and other cancers relative to normal bladder tissue, assays that evaluate the levels of 158P1D7 mRNA transcripts or proteins in a biological sample can be used to diagnose a disease associated with 158P1D7 dysregulation, and can provide prognostic information useful in defining appropriate therapeutic options.

The expression status of 158P1D7 provides information including the presence, stage and location of dysplastic, precancerous and cancerous cells, predicting susceptibility to various stages of disease, and/or for gauging tumor aggressiveness. Moreover, the expression profile makes it useful as an imaging reagent for metastasized disease. Consequently, an aspect of the invention is directed to the various molecular prognostic and diagnostic methods for examining the status of 158P1D7 in biological samples such as those from individuals suffering from, or suspected of suffering from a pathology characterized by dysregulated cellular growth, such as cancer.

As described above, the status of 158P1D7 in a biological sample can be examined by a number of well-known procedures in the art. For example, the status of 158P1D7 in a biological sample taken from a specific location in the body can be examined by evaluating the sample for the presence or absence of 158P1D7 expressing cells (e.g. those that express

158P1D7 mRNAs or proteins). This examination can provide evidence of dysregulated cellular growth, for example, when 158P1D7-expressing cells are found in a biological sample that does not normally contain such cells (such as a lymph node), because such alterations in the status of 158P1D7 in a biological sample are often associated with dysregulated cellular growth. Specifically, one indicator of dysregulated cellular growth is the metastases of cancer cells from an organ of origin (such as the bladder) to a different area of the body (such as a lymph node). By example, evidence of dysregulated cellular growth is important because occult lymph node metastases can be detected in a substantial proportion of patients with prostate cancer, and such metastases are associated with known predictors of disease progression (see, e.g., Murphy et al., Prostate 42(4): 315-317 (2000);Su et al., Semin. Surg. Oncol. 18(1): 17-28 (2000) and Freeman et al., J Urol 1995 Aug 154(2 Pt 1):474-8).

In one aspect, the invention provides methods for monitoring 158P1D7 gene products by determining the status of 158P1D7 gene products expressed by cells from an individual suspected of having a disease associated with dysregulated cell growth (such as hyperplasia or cancer) and then comparing the status so determined to the status of 158P1D7 gene products in a corresponding normal sample. The presence of aberrant 158P1D7 gene products in the test sample relative to the normal sample provides an indication of the presence of dysregulated cell growth within the cells of the individual.

In another aspect, the invention provides assays useful in determining the presence of cancer in an individual, comprising detecting a significant increase in 158P1D7 mRNA or protein expression in a test cell or tissue sample relative to expression levels in the corresponding normal cell or tissue. The presence of 158P1D7 mRNA can, for example, be evaluated in tissue samples including but not limited to those listed in Table I. The presence of significant 158P1D7 expression in any of these tissues is useful to indicate the emergence, presence and/or severity of a cancer, since the corresponding normal tissues do not express 158P1D7 mRNA or express it at lower levels.

In a related embodiment, 158P1D7 status is determined at the protein level rather than at the nucleic acid level. For example, such a method comprises determining the level of 158P1D7 protein expressed by cells in a test tissue sample and comparing the level so determined to the level of 158P1D7 expressed in a corresponding normal sample. In one embodiment, the presence of 158P1D7 protein is evaluated, for example, using immunohistochemical methods. 158P1D7 antibodies or binding partners capable of detecting 158P1D7 protein expression are used in a variety of assay formats well known in the art for this purpose.

In a further embodiment, one can evaluate the status of 158P1D7 nucleotide and amino acid sequences in a biological sample in order to identify perturbations in the structure of these molecules. These perturbations can include insertions, deletions, substitutions and the like. Such evaluations are useful because perturbations in the nucleotide and amino acid sequences are observed in a large number of proteins associated with a growth dysregulated phenotype (see, e.g., Marrogi et al., 1999, J. Cutan. Pathol. 26(8):369-378). For example, a mutation in the sequence of 158P1D7 may be indicative of the presence or promotion of a tumor. Such assays therefore have diagnostic and predictive value where a mutation in 158P1D7 indicates a potential loss of function or increase in tumor growth.

A wide variety of assays for observing perturbations in nucleotide and amino acid sequences are well known in the art. For example, the size and structure of nucleic acid or amino acid sequences of 158P1D7 gene products are observed by the Northern, Southern, Western, PCR and DNA sequencing protocols discussed herein. In addition, other methods for observing perturbations in nucleotide and amino acid sequences such as single strand conformation polymorphism analysis are well known in the art (see, e.g., U.S. Patent Nos. 5,382,510 issued 7 September 1999, and 5,952,170 issued 17 January 1995).

Additionally, one can examine the methylation status of the 158P1D7 gene in a biological sample. Aberrant demethylation and/or hypermethylation of CpG islands in gene 5' regulatory regions frequently occurs in immortalized and transformed cells, and can result in altered expression of various genes. For example, promoter hypermethylation of the DBCCR1, PAX6 and APC genes have been detected in bladder cancers leading to aberrant expression of the genes

(Esteller et al., Cancer Res 2001; 61:3225-3229) A variety of assays for examining methylation status of a gene are well known in the art. For example, one can utilize, in Southern hybridization approaches, methylation-sensitive restriction enzymes which cannot cleave sequences that contain methylated CpG sites to assess the methylation status of CpG islands. In addition, MSP (methylation specific PCR) can rapidly profile the methylation status of all the CpG sites present in a CpG island of a given gene. This procedure involves initial modification of DNA by sodium bisulfite (which will convert all unmethylated cytosines to uracil) followed by amplification using primers specific for methylated versus unmethylated DNA. Protocols involving methylation interference can also be found for example in Current Protocols In Molecular Biology, Unit 12, Frederick M. Ausubel et al. eds., 1995.

Gene amplification is an additional method for assessing the status of 158P1D7. Gene amplification is measured in a sample directly, for example, by conventional Southern blotting or Northern blotting to quantitate the transcription of mRNA (Thomas, 1980, Proc. Natl. Acad. Sci. USA, 77:5201-5205), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies are employed that recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn are labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Biopsied tissue or peripheral blood can be conveniently assayed for the presence of cancer cells using for example, Northern, dot blot or RT-PCR analysis to detect 158P1D7 expression. The presence of RT-PCR amplifiable 158P1D7 mRNA provides an indication of the presence of cancer. RT-PCR assays are well known in the art. RT-PCR detection assays for tumor cells in peripheral blood are currently being evaluated for use in the diagnosis and management of a number of human solid tumors.

A further aspect of the invention is an assessment of the susceptibility that an individual has for developing cancer. In one embodiment, a method for predicting susceptibility to cancer comprises detecting 158P1D7 mRNA or 158P1D7 protein in a tissue sample, its presence indicating susceptibility to cancer, wherein the degree of 158P1D7 mRNA expression correlates to the degree of susceptibility. In a specific embodiment, the presence of 158P1D7 in bladder or other tissue is examined, with the presence of 158P1D7 in the sample providing an indication of bladder cancer susceptibility (or the emergence or existence of a bladder tumor). Similarly, one can evaluate the integrity 158P1D7 nucleotide and amino acid sequences in a biological sample, in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like. The presence of one or more perturbations in 158P1D7 gene products in the sample is an indication of cancer susceptibility (or the emergence or existence of a tumor).

The invention also comprises methods for gauging tumor aggressiveness. In one embodiment, a method for gauging aggressiveness of a tumor comprises determining the level of 158P1D7 mRNA or 158P1D7 protein expressed by tumor cells, comparing the level so determined to the level of 158P1D7 mRNA or 158P1D7 protein expressed in a corresponding normal tissue taken from the same individual or a normal tissue reference sample, wherein the degree of 158P1D7 mRNA or 158P1D7 protein expression in the tumor sample relative to the normal sample indicates the degree of aggressiveness. In a specific embodiment, aggressiveness of a tumor is evaluated by determining the extent to which 158P1D7 is expressed in the tumor cells, with higher expression levels indicating more aggressive tumors. Another embodiment is the evaluation of the integrity of 158P1D7 nucleotide and amino acid sequences in a biological sample, in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like. The presence of one or more perturbations indicates more aggressive tumors.

Another embodiment of the invention is directed to methods for observing the progression of a malignancy in an individual over time. In one embodiment, methods for observing the progression of a malignancy in an individual over time comprise determining the level of 158P1D7 mRNA or 158P1D7 protein expressed by cells in a sample of the tumor, comparing

the level so determined to the level of 158P1D7 mRNA or 158P1D7 protein expressed in an equivalent tissue sample taken from the same individual at a different time, wherein the degree of 158P1D7 mRNA or 158P1D7 protein expression in the tumor sample over time provides information on the progression of the cancer. In a specific embodiment, the progression of a cancer is evaluated by determining 158P1D7 expression in the tumor cells over time, where increased expression over time indicates a progression of the cancer. Also, one can evaluate the integrity 158P1D7 nucleotide and amino acid sequences in a biological sample in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like, where the presence of one or more perturbations indicates a progression of the cancer.

The above diagnostic approaches can be combined with any one of a wide variety of prognostic and diagnostic protocols known in the art. For example, another embodiment of the invention is directed to methods for observing a coincidence between the expression of 158P1D7 gene and 158P1D7 gene products (or perturbations in 158P1D7 gene and 158P1D7 gene products) and a factor that is associated with malignancy, as a means for diagnosing and prognosticating the status of a tissue sample. A wide variety of factors associated with malignancy can be utilized, such as the expression of genes associated with malignancy (e.g. PSCA, H-rasand p53 expression etc.) as well as gross cytological observations (see, e.g., Bocking et al., 1984, Anal. Quant. Cytol. 6(2):74-88; Epstein, 1995, Hum. Pathol. 26(2):223-9; Thorson et al., 1998, Mod. Pathol. 11(6):543-51; Baisden et al., 1999, Am. J. Surg. Pathol. 23(8):918-24). Methods for observing a coincidence between the expression of 158P1D7 gene and 158P1D7 gene products (or perturbations in 158P1D7 gene and 158P1D7 gene products) and another factor that is associated with malignancy are useful, for example, because the presence of a set of specific factors that coincide with disease provides information crucial for diagnosing and prognosticating the status of a tissue sample.

In one embodiment, methods for observing a coincidence between the expression of 158P1D7 gene and 158P1D7 gene products (or perturbations in 158P1D7 gene and 158P1D7 gene products) and another factor associated with malignancy entails detecting the overexpression of 158P1D7 mRNA or protein in a tissue sample, detecting the overexpression of BLCA-4A mRNA or protein in a tissue sample (or PSCA expression), and observing a coincidence of 158P1D7 mRNA or protein and BLCA-4 mRNA or protein overexpression (or PSCA expression) (Amara et al., 2001, Cancer Res 61:4660-4665; Konety et al., Clin Cancer Res, 2000, 6(7):2618-2625). In a specific embodiment, the expression of 158P1D7 and BLCA-4 mRNA in bladder tissue is examined, where the coincidence of 158P1D7 and BLCA-4 mRNA overexpression in the sample indicates the existence of bladder cancer, bladder cancer susceptibility or the emergence or status of a bladder tumor.

Methods for detecting and quantifying the expression of 158P1D7 mRNA or protein are described herein, and standard nucleic acid and protein detection and quantification technologies are well known in the art. Standard methods for the detection and quantification of 158P1D7 mRNA include *in situ* hybridization using labeled 158P1D7 riboprobes, Northern blot and related techniques using 158P1D7 polynucleotide probes, RT-PCR analysis using primers specific for 158P1D7, and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like. In a specific embodiment, semi-quantitative RT-PCR is used to detect and quantify 158P1D7 mRNA expression. Any number of primers capable of amplifying 158P1D7 can be used for this purpose, including but not limited to the various primer sets specifically described herein. In a specific embodiment, polyclonal or monoclonal antibodies specifically reactive with the wild-type 158P1D7 protein can be used in an immunohistochemical assay of biopsied tissue.

IX.) Identification of Molecules That Interact With 158P1D7

The 158P1D7 protein and nucleic acid sequences disclosed herein allow a skilled artisan to identify proteins, small molecules and other agents that interact with 158P1D7, as well as pathways activated by 158P1D7 via any one of a variety of art accepted protocols. For example, one can utilize one of the so-called interaction trap systems (also referred to as the "two-hybrid assay"). In such systems, molecules interact and reconstitute a transcription factor which directs expression of a reporter gene, whereupon the expression of the reporter gene is assayed. Other systems identify protein-protein interactions

in vivo through reconstitution of a eukaryotic transcriptional activator, see, e.g., U.S. Patent Nos. 5,955,280 issued 21 September 1999, 5,925,523 issued 20 July 1999, 5,846,722 issued 8 December 1998 and 6,004,746 issued 21 December 1999. Algorithms are also available in the art for genome-based predictions of protein function (see, e.g., Marcotte, et al., Nature 402: 4 November 1999, 83-86).

Alternatively one can screen peptide libraries to identify molecules that interact with 158P1D7 protein sequences. In such methods, peptides that bind to 158P1D7 are identified by screening libraries that encode a random or controlled collection of amino acids. Peptides encoded by the libraries are expressed as fusion proteins of bacteriophage coat proteins, the bacteriophage particles are then screened against the 158P1D7 protein.

Accordingly, peptides having a wide variety of uses, such as therapeutic, prognostic or diagnostic reagents, are thus identified without any prior information on the structure of the expected ligand or receptor molecule. Typical peptide libraries and screening methods that can be used to identify molecules that interact with 158P1D7 protein sequences are disclosed for example in U.S. Patent Nos. 5,723,286 issued 3 March 1998 and 5,733,731 issued 31 March 1998.

Alternatively, cell lines that express 158P1D7 are used to identify protein-protein interactions mediated by 158P1D7. Such interactions can be examined using immunoprecipitation techniques (see, e.g., Hamilton BJ, et al. Biochem. Biophys. Res. Commun. 1999, 261:646-51). 158P1D7 protein can be immunoprecipitated from 158P1D7-expressing cell lines using anti-158P1D7 antibodies. Alternatively, antibodies against His-tag can be used in a cell line engineered to express fusions of 158P1D7 and a His-tag (vectors mentioned above). The immunoprecipitated complex can be examined for protein association by procedures such as Western blotting, ^{35}S-methionine labeling of proteins, protein microsequencing, silver staining and two-dimensional gel electrophoresis.

Small molecules and ligands that interact with 158P1D7 can be identified through related embodiments of such screening assays. For example, small molecules can be identified that interfere with protein function, including molecules that interfere with 158P1D7's ability to mediate phosphorylation and de-phosphorylation, interaction with DNA or RNA molecules as an indication of regulation of cell cycles, second messenger signaling or tumorigenesis. Similarly, small molecules that modulate 158P1D7 related ion channel, protein pump, or cell communication functions 158P1D7are identified and used to treat patients that have a cancer that expresses 158P1D7 (see, e.g., Hille, B., Ionic Channels of Excitable Membranes 2nd Ed., Sinauer Assoc., Sunderland, MA, 1992). Moreover, ligands that regulate 158P1D7 function can be identified based on their ability to bind 158P1D7 and activate a reporter construct. Typical methods are discussed for example in U.S. Patent No. 5,928,868 issued 27 July 1999, and include methods for forming hybrid ligands in which at least one ligand is a small molecule. In an illustrative embodiment, cells engineered to express a fusion protein of 158P1D7 and a DNA-binding protein are used to co-express a fusion protein of a hybrid ligand/small molecule and a cDNA library transcriptional activator protein. The cells further contain a reporter gene, the expression of which is conditioned on the proximity of the first and second fusion proteins to each other, an event that occurs only if the hybrid ligand binds to target sites on both hybrid proteins. Those cells that express the reporter gene are selected and the unknown small molecule or the unknown ligand is identified. This method provides a means of identifying modulators which activate or inhibit 158P1D7.

An embodiment of this invention comprises a method of screening for a molecule that interacts with an 158P1D7 amino acid sequence shown in Figure 2 or Figure 3, comprising the steps of contacting a population of molecules with the 158P1D7 amino acid sequence, allowing the population of molecules and the 158P1D7 amino acid sequence to interact under conditions that facilitate an interaction, determining the presence of a molecule that interacts with the 158P1D7 amino acid sequence, and then separating molecules that do not interact with the 158P1D7 amino acid sequence from molecules that do. In a specific embodiment, the method further comprises purifying, characterizing and identifying a molecule that interacts with the 158P1D7 amino acid sequence. The identified molecule can be used to modulate a function performed by 158P1D7. In a preferred embodiment, the 158P1D7 amino acid sequence is contacted with a library of peptides.

X.) Therapeutic Methods and Compositions

The identification of 158P1D7 as a protein that is normally expressed in a restricted set of tissues, but which is also expressed in bladder and other cancers, opens a number of therapeutic approaches to the treatment of such cancers. As contemplated herein, 158P1D7 functions as a transcription factor involved in activating tumor-promoting genes or repressing genes that block tumorigenesis.

Accordingly, therapeutic approaches that inhibit the activity of the 158P1D7 protein are useful for patients suffering from a cancer that expresses 158P1D7. These therapeutic approaches generally fall into two classes. One class comprises various methods for inhibiting the binding or association of the 158P1D7 protein with its binding partner or with other proteins. Another class comprises a variety of methods for inhibiting the transcription of the 158P1D7 gene or translation of 158P1D7 mRNA.

X.A.) Anti-Cancer Vaccines

The invention provides cancer vaccines comprising a 158P1D7-related protein or 158P1D7-related nucleic acid. In view of the expression of 158P1D7, cancer vaccines prevent and/or treat 158P1D7-expressing cancers with minimal or no effects on non-target tissues. The use of a tumor antigen in a vaccine that generates humoral and/or cell-mediated immune responses as anti-cancer therapy is well known in the art (see, e.g., Hodge et al., 1995, Int. J. Cancer 63:231-237; Fong et al., 1997, J. Immunol. 159:3113-3117).

Such methods can be readily practiced by employing a 158P1D7-related protein, or a 158P1D7-encoding nucleic acid molecule and recombinant vectors capable of expressing and presenting the 158P1D7 immunogen (which typically comprises a number of antibody or T cell epitopes). Skilled artisans understand that a wide variety of vaccine systems for delivery of immunoreactive epitopes are known in the art (see, e.g., Heryln et al., Ann Med 1999 Feb 31(1):66-78; Maruyama et al., Cancer Immunol Immunother 2000 Jun 49(3):123-32) Briefly, such methods of generating an immune response (e.g. humoral and/or cell-mediated) in a mammal, comprise the steps of: exposing the mammal's immune system to an immunoreactive epitope (e.g. an epitope present in the 158P1D7 protein shown in Figure 2 or analog or homolog thereof) so that the mammal generates an immune response that is specific for that epitope (e.g. generates antibodies that specifically recognize that epitope). In a preferred method, the 158P1D7 immunogen contains a biological motif, see e.g., Tables V-XVIII, or a peptide of a size range from 158P1D7 indicated in Figure 11, Figure 12, Figure 13, Figure 14, and Figure 15.

The entire 158P1D7 protein, immunogenic regions or epitopes thereof can be combined and delivered by various means. Such vaccine compositions can include, for example, lipopeptides (*e.g.*,Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991: Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995), peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L.,

Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

In patients with 158P1D7-associated cancer, the vaccine compositions of the invention can also be used in conjunction with other treatments used for cancer, *e.g.*, surgery, chemotherapy, drug therapies, radiation therapies, *etc.* including use in combination with immune adjuvants such as IL-2, IL-12, GM-CSF, and the like.

Cellular Vaccines:

CTL epitopes can be determined using specific algorithms to identify peptides within 158P1D7 protein that bind corresponding HLA alleles (see e.g., Table IV; Epimer™ and Epimatrix™, Brown University (URL brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html); and, BIMAS, (URL bimas.dcrt.nih.gov/; SYFPEITHI at URL syfpeithi.bmi-heidelberg.com/). In a preferred embodiment, the 158P1D7 immunogen contains one or more amino acid sequences identified using techniques well known in the art, such as the sequences shown in Tables V-XVIII or a peptide of 8, 9, 10 or 11 amino acids specified by an HLA Class I motif/supermotif (e.g., Table IV (A), Table IV (D), or Table IV (E)) and/or a peptide of at least 9 amino acids that comprises an HLA Class II motif/supermotif (e.g., Table IV (B) or Table IV (C)). As is appreciated in the art, the HLA Class I binding groove is essentially closed ended so that peptides of only a particular size range can fit into the groove and be bound, generally HLA Class I epitopes are 8, 9, 10, or 11 amino acids long. In contrast, the HLA Class II binding groove is essentially open ended; therefore a peptide of about 9 or more amino acids can be bound by an HLA Class II molecule. Due to the binding groove differences between HLA Class I and II, HLA Class I motifs are length specific, i.e., position two of a Class I motif is the second amino acid in an amino to carboxyl direction of the peptide. The amino acid positions in a Class II motif are relative only to each other, not the overall peptide, i.e., additional amino acids can be attached to the amino and/or carboxyl termini of a motif-bearing sequence. HLA Class II epitopes are often 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids long, or longer than 25 amino acids.

Antibody-based Vaccines

A wide variety of methods for generating an immune response in a mammal are known in the art (for example as the first step in the generation of hybridomas). Methods of generating an immune response in a mammal comprise exposing the mammal's immune system to an immunogenic epitope on a protein (e.g. the 158P1D7 protein) so that an immune response is generated. A typical embodiment consists of a method for generating an immune response to 158P1D7 in a host, by contacting the host with a sufficient amount of at least one 158P1D7 B cell or cytotoxic T-cell epitope or analog thereof; and at least one periodic interval thereafter re-contacting the host with the 158P1D7 B cell or cytotoxic T-cell epitope or analog thereof. A specific embodiment consists of a method of generating an immune response against a 158P1D7-related protein or a man-made multiepitopic peptide comprising: administering 158P1D7 immunogen (e.g. the 158P1D7 protein or a peptide fragment thereof, an 158P1D7 fusion protein or analog etc.) in a vaccine preparation to a human or another mammal. Typically, such vaccine preparations further contain a suitable adjuvant (see, e.g., U.S. Patent No. 6,146,635) or a universal helper epitope such as a PADRE™ peptide (Epimmune Inc., San Diego, CA; see, e.g., Alexander et al., J. Immunol. 2000 164(3); 164(3): 1625-1633; Alexander et al., Immunity 1994 1(9): 751-761 and Alexander et al., Immunol. Res. 1998 18(2): 79-92). An alternative method comprises generating an immune response in an individual against a 158P1D7 immunogen by: administering *in vivo* to muscle or skin of the individual's body a DNA molecule that comprises a DNA sequence that encodes an 158P1D7 immunogen, the DNA sequence operatively linked to regulatory sequences which control the expression of the DNA sequence; wherein the DNA molecule is taken up by cells, the DNA sequence is expressed in the cells and an immune response is generated against the immunogen (see, e.g., U.S. Patent No.

5,962,428). Optionally a genetic vaccine facilitator such as anionic lipids; saponins; lectins; estrogenic compounds; hydroxylated lower alkyls; dimethyl sulfoxide; and urea is also administered.

Nucleic Acid Vaccines:

Vaccine compositions of the invention include nucleic acid-mediated modalities. DNA or RNA that encode protein(s) of the invention can be administered to a patient. Genetic immunization methods can be employed to generate prophylactic or therapeutic humoral and cellular immune responses directed against cancer cells expressing 158P1D7. Constructs comprising DNA encoding a 158P1D7-related protein/immunogen and appropriate regulatory sequences can be injected directly into muscle or skin of an individual, such that the cells of the muscle or skin take-up the construct and express the encoded 158P1D7 protein/immunogen. Alternatively, a vaccine comprises a 158P1D7-related protein. Expression of the 158P1D7-related protein immunogen results in the generation of prophylactic or therapeutic humoral and cellular immunity against cells that bear 158P1D7 protein. Various prophylactic and therapeutic genetic immunization techniques known in the art can be used (for review, see information and references published at Internet address URL: genweb.com). Nucleic acid-based delivery is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, proteins of the invention can be expressed via viral or bacterial vectors. Various viral gene delivery systems that can be used in the practice of the invention include, but are not limited to, vaccinia, fowlpox, canarypox, adenovirus, influenza, poliovirus, adeno-associated virus, lentivirus, and sindbis virus (see, e.g., Restifo, 1996, Curr. Opin. Immunol. 8:658-663; Tsang et al. J. Natl. Cancer Inst. 87:982-990 (1995)). Non-viral delivery systems can also be employed by introducing naked DNA encoding a 158P1D7-related protein into the patient (e.g., intramuscularly or intradermally) to induce an anti-tumor response.

Vaccinia virus is used, for example, as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the protein immunogenic peptide, and thereby elicits a host immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

Thus, gene delivery systems are used to deliver a 158P1D7-related nucleic acid molecule. In one embodiment, the full-length human 158P1D7 cDNA is employed. In another embodiment, 158P1D7 nucleic acid molecules encoding specific cytotoxic T lymphocyte (CTL) and/or antibody epitopes are employed.

Ex Vivo Vaccines

Various *ex vivo* strategies can also be employed to generate an immune response. One approach involves the use of antigen presenting cells (APCs) such as dendritic cells (DC) to present 158P1D7 antigen to a patient's immune system. Dendritic cells express MHC class I and II molecules, B7 co-stimulator, and IL-12, and are thus highly specialized antigen presenting cells. In bladder cancer, autologous dendritic cells pulsed with peptides of the MAGE-3 antigen are being used in a Phase I clinical trial to stimulate bladder cancer patients' immune systems (Nishiyama et al., 2001, Clin Cancer Res, 7(1):23-31). Thus, dendritic cells can be used to present 158P1D7 peptides to T cells in the context of MHC class I or II molecules. In one embodiment, autologous dendritic cells are pulsed with 158P1D7 peptides capable of binding to MHC class I and/or class II molecules. In another embodiment, dendritic cells are pulsed with the complete 158P1D7 protein. Yet another embodiment involves engineering the overexpression of the 158P1D7 gene in dendritic cells using various implementing vectors known in

the art, such as adenovirus (Arthur et al., 1997, Cancer Gene Ther. 4:17-25), retrovirus (Henderson et al., 1996, Cancer Res. 56:3763-3770), lentivirus, adeno-associated virus, DNA transfection (Ribas et al., 1997, Cancer Res. 57:2865-2869), or tumor-derived RNA transfection (Ashley et al., 1997, J. Exp. Med. 186:1177-1182). Cells that express 158P1D7 can also be engineered to express immune modulators, such as GM-CSF, and used as immunizing agents.

X.B.) 158P1D7 as a Target for Antibody-based Therapy

158P1D7 is an attractive target for antibody-based therapeutic strategies. A number of antibody strategies are known in the art for targeting both extracellular and intracellular molecules (see, e.g., complement and ADCC mediated killing as well as the use of intrabodies). Because 158P1D7 is expressed by cancer cells of various lineages relative to corresponding normal cells, systemic administration of 158P1D7-immunoreactive compositions are prepared that exhibit excellent sensitivity without toxic, non-specific and/or non-target effects caused by binding of the immunoreactive composition to non-target organs and tissues. Antibodies specifically reactive with domains of 158P1D7 are useful to treat 158P1D7-expressing cancers systemically, either as conjugates with a toxin or therapeutic agent, or as naked antibodies capable of inhibiting cell proliferation or function.

158P1D7 antibodies can be introduced into a patient such that the antibody binds to 158P1D7 and modulates a function, such as an interaction with a binding partner, and consequently mediates destruction of the tumor cells and/or inhibits the growth of the tumor cells. Mechanisms by which such antibodies exert a therapeutic effect can include complement-mediated cytolysis, antibody-dependent cellular cytotoxicity, modulation of the physiological function of 158P1D7, inhibition of ligand binding or signal transduction pathways, modulation of tumor cell differentiation, alteration of tumor angiogenesis factor profiles, and/or apoptosis.

Those skilled in the art understand that antibodies can be used to specifically target and bind immunogenic molecules such as an immunogenic region of the 158P1D7 sequence shown in Figure 2 or Figure 3. In addition, skilled artisans understand that it is routine to conjugate antibodies to cytotoxic agents (see, e.g., Slevers et al. Blood 93:11 3678-3684 (June 1, 1999)). When cytotoxic and/or therapeutic agents are delivered directly to cells, such as by conjugating them to antibodies specific for a molecule expressed by that cell (e.g. 158P1D7), the cytotoxic agent will exert its known biological effect (i.e. cytotoxicity) on those cells.

A wide variety of compositions and methods for using antibody-cytotoxic agent conjugates to kill cells are known in the art. In the context of cancers, typical methods entail administering to an animal having a tumor a biologically effective amount of a conjugate comprising a selected cytotoxic and/or therapeutic agent linked to a targeting agent (e.g. an anti-158P1D7 antibody) that binds to a marker (e.g. 158P1D7) expressed, accessible to binding or localized on the cell surfaces. A typical embodiment is a method of delivering a cytotoxic and/or therapeutic agent to a cell expressing 158P1D7, comprising conjugating the cytotoxic agent to an antibody that immunospecifically binds to a 158P1D7 epitope, and, exposing the cell to the antibody-agent conjugate. Another illustrative embodiment is a method of treating an individual suspected of suffering from metastasized cancer, comprising a step of administering parenterally to said individual a pharmaceutical composition comprising a therapeutically effective amount of an antibody conjugated to a cytotoxic and/or therapeutic agent.

Cancer immunotherapy using anti-158P1D7 antibodies can be done in accordance with various approaches that have been successfully employed in the treatment of other types of cancer, including but not limited to colon cancer (Arlen et al., 1998, Crit. Rev. Immunol. 18:133-138), multiple myeloma (Ozaki et al., 1997, Blood 90:3179-3186, Tsunenari et al., 1997, Blood 90:2437-2444), gastric cancer (Kasprzyk et al., 1992, Cancer Res. 52:2771-2776), B-cell lymphoma (Funakoshi et al., 1996, J. Immunother. Emphasis Tumor Immunol. 19:93-101), leukemia (Zhong et al., 1996, Leuk. Res. 20:581-589), colorectal cancer (Moun et al., 1994, Cancer Res. 54:6160-6166; Velders et al., 1995, Cancer Res. 55:4398-4403), and breast cancer (Shepard et al., 1991, J. Clin. Immunol. 11:117-127). Some therapeutic approaches involve conjugation of

naked antibody to a toxin, such as the conjugation of Y^{91} or I^{131} to anti-CD20 antibodies (e.g., Zevalin™, IDEC Pharmaceuticals Corp. or Bexxar™, Coulter Pharmaceuticals), while others involve co-administration of antibodies and other therapeutic agents, such as Herceptin™ (trastuzumab) with paclitaxel (Genentech, Inc.). To treat bladder cancer, for example, 158P1D7 antibodies can be administered in conjunction with radiation, chemotherapy or hormone ablation.

Although 158P1D7 antibody therapy is useful for all stages of cancer, antibody therapy can be particularly appropriate in advanced or metastatic cancers. Treatment with the antibody therapy of the invention is indicated for patients who have received one or more rounds of chemotherapy. Alternatively, antibody therapy of the invention is combined with a chemotherapeutic or radiation regimen for patients who have not received chemotherapeutic treatment. Additionally, antibody therapy can enable the use of reduced dosages of concomitant chemotherapy, particularly for patients who do not tolerate the toxicity of the chemotherapeutic agent very well.

Cancer patients can be evaluated for the presence and level of 158P1D7 expression, preferably using immunohistochemical assessments of tumor tissue, quantitative 158P1D7 imaging, or other techniques that reliably indicate the presence and degree of 158P1D7 expression. Immunohistochemical analysis of tumor biopsies or surgical specimens is preferred for this purpose. Methods for immunohistochemical analysis of tumor tissues are well known in the art.

Anti-158P1D7 monoclonal antibodies that treat bladder and other cancers include those that initiate a potent immune response against the tumor or those that are directly cytotoxic. In this regard, anti-158P1D7 monoclonal antibodies (mAbs) can elicit tumor cell lysis by either complement-mediated or antibody-dependent cell cytotoxicity (ADCC) mechanisms, both of which require an intact Fc portion of the immunoglobulin molecule for interaction with effector cell Fc receptor sites on complement proteins. In addition, anti-158P1D7 mAbs that exert a direct biological effect on tumor growth are useful to treat cancers that express 158P1D7. Mechanisms by which directly cytotoxic mAbs act include: inhibition of cell growth, modulation of cellular differentiation, modulation of tumor angiogenesis factor profiles, and the induction of apoptosis. The mechanism(s) by which a particular anti-158P1D7 mAb exerts an anti-tumor effect is evaluated using any number of *in vitro* assays that evaluate cell death such as ADCC, ADMMC, complement-mediated cell lysis, and so forth, as is generally known in the art.

In some patients, the use of murine or other non-human monoclonal antibodies, or human/mouse chimeric mAbs can induce moderate to strong immune responses against the non-human antibody. This can result in clearance of the antibody from circulation and reduced efficacy. In the most severe cases, such an immune response can lead to the extensive formation of immune complexes which, potentially, can cause renal failure. Accordingly, preferred monoclonal antibodies used in the therapeutic methods of the invention are those that are either fully human or humanized and that bind specifically to the target 158P1D7 antigen with high affinity but exhibit low or no antigenicity in the patient.

Therapeutic methods of the invention contemplate the administration of single anti-158P1D7 mAbs as well as combinations, or cocktails, of different mAbs. Such mAb cocktails can have certain advantages inasmuch as they contain mAbs that target different epitopes, exploit different effector mechanisms or combine directly cytotoxic mAbs with mAbs that rely on immune effector functionality. Such mAbs in combination can exhibit synergistic therapeutic effects. In addition, anti-158P1D7 mAbs can be administered concomitantly with other therapeutic modalities, including but not limited to various chemotherapeutic agents, androgen-blockers, immune modulators (e.g., IL-2, GM-CSF), surgery or radiation. The anti-158P1D7 mAbs are administered in their "naked" or unconjugated form, or can have a therapeutic agent(s) conjugated to them.

Anti-158P1D7 antibody formulations are administered via any route capable of delivering the antibodies to a tumor cell. Routes of administration include, but are not limited to, intravenous, intraperitoneal, intramuscular, intratumor, intradermal, and the like. Treatment generally involves repeated administration of the anti-158P1D7 antibody preparation,

via an acceptable route of administration such as intravenous injection (IV), typically at a dose in the range of about 0.1 to about 10 mg/kg body weight. In general, doses in the range of 10-500 mg mAb per week are effective and well tolerated.

Based on clinical experience with the Herceptin mAb in the treatment of metastatic breast cancer, an initial loading dose of approximately 4 mg/kg patient body weight IV, followed by weekly doses of about 2 mg/kg IV of the anti- 158P1D7 mAb preparation represents an acceptable dosing regimen. Preferably, the initial loading dose is administered as a 90 minute or longer infusion. The periodic maintenance dose is administered as a 30 minute or longer infusion, provided the initial dose was well tolerated. As appreciated by those of skill in the art, various factors can influence the ideal dose regimen in a particular case. Such factors include, for example, the binding affinity and half life of the Ab or mAbs used, the degree of 158P1D7 expression in the patient, the extent of circulating shed 158P1D7 antigen, the desired steady-state antibody concentration level, frequency of treatment, and the influence of chemotherapeutic or other agents used in combination with the treatment method of the invention, as well as the health status of a particular patient.

Optionally, patients should be evaluated for the levels of 158P1D7 in a given sample (e.g. the levels of circulating 158P1D7 antigen and/or 158P1D7 expressing cells) in order to assist in the determination of the most effective dosing regimen, etc. Such evaluations are also used for monitoring purposes throughout therapy, and are useful to gauge therapeutic success in combination with the evaluation of other parameters (for example, urine cytology and/or ImmunoCyt levels in bladder cancer therapy, or by analogy, serum PSA levels in prostate cancer therapy).

Anti-idiotypic anti-158P1D7 antibodies can also be used in anti-cancer therapy as a vaccine for inducing an immune response to cells expressing a 158P1D7-related protein. In particular, the generation of anti-idiotypic antibodies is well known in the art; this methodology can readily be adapted to generate anti-idiotypic anti-158P1D7 antibodies that mimic an epitope on a 158P1D7-related protein (see, for example, Wagner et al., 1997, Hybridoma 16: 33-40; Foon et al., 1995, J. Clin. Invest. 96:334-342; Herlyn et al., 1996, Cancer Immunol. Immunother. 43:65-76). Such an anti-idiotypic antibody can be used in cancer vaccine strategies.

X.C.) 158P1D7 as a Target for Cellular Immune Responses

Vaccines and methods of preparing vaccines that contain an immunogenically effective amount of one or more HLA-binding peptides as described herein are further embodiments of the invention. Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptides. A peptide can be present in a vaccine individually. Alternatively, the peptide can exist as a homopolymer comprising multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased immunological reaction and, where different peptide epitopes are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targeted for an immune response. The composition can be a naturally occurring region of an antigen or can be prepared, *e.g.*, recombinantly or by chemical synthesis.

Carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza, hepatitis B virus core protein, and the like. The vaccines can contain a physiologically tolerable (*i.e.*, acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerylcysteinlyseryl- serine (P_3CSS). Moreover, an adjuvant such as a synthetic cytosine-phosphorothiolated-guanine-containing (CpG) oligonucleotides has been found to increase CTL responses 10- to 100-fold. (see, e.g. Davila and Celis J. Immunol. 165:539-547 (2000))

Upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs and/or HTLs specific for the desired antigen. Consequently, the host becomes at least partially immune to later development of cells that express or overexpress 158P1D7 antigen, or derives at least some therapeutic benefit when the antigen was tumor-associated.

In some embodiments, it may be desirable to combine the class I peptide components with components that induce or facilitate neutralizing antibody and or helper T cell responses directed to the target antigen. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a cross reactive HTL epitope such as PADRE™ (Epimmune, San Diego, CA) molecule (described *e.g.*, in U.S. Patent Number 5,736,142).

A vaccine of the invention can also include antigen-presenting cells (APC), such as dendritic cells (DC), as a vehicle to present peptides of the invention. Vaccine compositions can be created *in vitro*, following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs *in vitro*. For example, dendritic cells are transfected, *e.g.*, with a minigene in accordance with the invention, or are pulsed with peptides. The dendritic cell can then be administered to a patient to elicit immune responses *in vivo*. Vaccine compositions, either DNA- or peptide-based, can also be administered *in vivo* in combination with dendritic cell mobilization whereby loading of dendritic cells occurs *in vivo*.

Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. It is preferred that each of the following principles be balanced in order to make the selection. The multiple epitopes to be incorporated in a given vaccine composition may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived.

1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For HLA Class I this includes 3-4 epitopes that come from at least one tumor associated antigen (TAA). For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one TAA (*see, e.g.*, Rosenberg *et al.*, *Science* 278:1447-1450). Epitopes from one TAA may be used in combination with epitopes from one or more additional TAAs to produce a vaccine that targets tumors with varying expression patterns of frequently-expressed TAAs.

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, often 200 nM or less; and for Class II an IC_{50} of 1000 nM or less.

3.) Sufficient supermotif bearing-peptides, or a sufficient array of allele-specific motif-bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth, or redundancy of, population coverage.

4.) When selecting epitopes from cancer-related antigens it is often useful to select analogs because the patient may have developed tolerance to the native epitope.

5.) Of particular relevance are epitopes referred to as "nested epitopes." Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A nested peptide sequence can comprise B cell, HLA class I and/or HLA class II epitopes. When providing nested epitopes, a general objective is to provide the greatest number of epitopes per sequence. Thus, an aspect is to avoid providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a multi-epitopic sequence,

such as a sequence comprising nested epitopes, it is generally important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

6.) If a polyepitopic protein is created, or when creating a minigene, an objective is to generate the smallest peptide that encompasses the epitopes of interest. This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial polyepitopic peptide, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can, for example, be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

7.) Where the sequences of multiple variants of the same target protein are present, potential peptide epitopes can also be selected on the basis of their conservancy. For example, a criterion for conservancy may define that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide be conserved in a designated percentage of the sequences evaluated for a specific protein antigen.

X.C.1. Minigene Vaccines

A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a peptide comprising one or multiple epitopes of the invention.

The use of multi-epitope minigenes is described below and in, Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid encoding supermotif- and/or motif-bearing epitopes derived 158P1D7, the PADRE® universal helper T cell epitope (or multiple HTL epitopes from 158P1D7), and an endoplasmic reticulum-translocating signal sequence can be engineered. A vaccine may also comprise epitopes that are derived from other TAAs.

The immunogenicity of a multi-epitopic minigene can be confirmed in transgenic mice to evaluate the magnitude of CTL induction responses against the epitopes tested. Further, the immunogenicity of DNA-encoded epitopes *in vivo* can be correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these experiments can show that the minigene serves to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous polypeptide sequence is created. To optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that can be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, antibody epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or for HTL responses, pan-DR binding proteins (PADRE™, Epimmune, San Diego, CA). Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF-β) may be beneficial in certain diseases.

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor according to well-known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for formulating purified plasmid DNA may be desirable. A variety of methods

have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, as described by WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Pat No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and HLA class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 (^{51}Cr) labeled and used as target cells for epitope-specific CTL lines; cytolysis, detected by ^{51}Cr release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (i.p.) for lipid-complexed DNA). Twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded, ^{51}Cr-labeled target cells using standard techniques. Lysis of target cells that were sensitized by HLA loaded with peptide epitopes, corresponding to minigene-encoded epitopes, demonstrates DNA vaccine function for *in vivo* induction of CTLs. Immunogenicity of HTL epitopes is confirmed in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

Minigenes can also be delivered using other bacterial or viral delivery systems well known in the art, *e.g.*, an expression construct encoding epitopes of the invention can be incorporated into a viral vector such as vaccinia.

X.C.2. Combinations of CTL Peptides with Helper Peptides

Vaccine compositions comprising CTL peptides of the invention can be modified, e.g., analoged, to provide desired attributes, such as improved serum half life, broadened population coverage or enhanced immunogenicity.

For instance, the ability of a peptide to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. Although a CTL peptide can be directly linked to a T helper peptide, often CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues and sometimes 10 or more residues. The CTL peptide epitope can be linked to the T helper peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in a majority of a genetically diverse population. This can be accomplished by selecting peptides that bind to many, most, or all of the HLA class II molecules. Examples of such amino acid bind many HLA Class II molecules include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE; SEQ ID NO: 24), *Plasmodium falciparum* circumsporozoite (CS) protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS; SEQ ID NO: 25), and *Streptococcus* 18kD protein at positions 116-131 (GAVDSILGGVATYGAA; SEQ ID NO: 26). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (*see, e.g.*, PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (*e.g.*, PADRE™, Epimmune, Inc., San Diego, CA) are designed to most preferably bind most HLA-DR (human HLA class II) molecules. For instance, a pan-DR-binding epitope peptide having the formula: aKXVAAWTLKAAa (SEQ ID NO: 27), where "X" is either cyclohexylalanine, phenylalanine, or tyrosine, and a is either D-alanine or L-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type. An alternative of a pan-DR binding epitope comprises all "L" natural amino acids and can be provided in the form of nucleic acids that encode the epitope.

HTL peptide epitopes can also be modified to alter their biological properties. For example, they can be modified to include D-amino acids to increase their resistance to proteases and thus extend their serum half life, or they can be conjugated to other molecules such as lipids, proteins, carbohydrates, and the like to increase their biological activity. For example, a T helper peptide can be conjugated to one or more palmitic acid chains at either the amino or carboxyl termini.

X.C.3. Combinations of CTL Peptides with T Cell Priming Agents

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes B lymphocytes or T lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo*. For example, palmitic acid residues can be attached to the ε-and α- amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. In a preferred embodiment, a particularly effective immunogenic composition comprises palmitic acid attached to ε- and α- amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinlyseryl- serine (P_3CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide (*see, e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Peptides of the invention can be coupled to P_3CSS, for example, and the lipopeptide administered to an individual to specifically prime an immune response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P_3CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses.

X.C.4. Vaccine Compositions Comprising DC Pulsed with CTL and/or HTL Peptides

An embodiment of a vaccine composition in accordance with the invention comprises *ex vivo* administration of a cocktail of epitope-bearing peptides to PBMC, or isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Pharmacia-Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes complexed with HLA molecules on their surfaces.

The DC can be pulsed *ex vivo* with a cocktail of peptides, some of which stimulate CTL responses to 158P1D7. Optionally, a helper T cell (HTL) peptide, such as a natural or artificial loosely restricted HLA Class II peptide, can be included to facilitate the CTL response. Thus, a vaccine in accordance with the invention is used to treat a cancer which expresses or overexpresses 158P1D7.

X.D. Adoptive Immunotherapy

Antigenic 158P1D7-related peptides are used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTL or HTL cells, can be used to treat tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular antigen are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (e.g., a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells.

X.E. Administration of Vaccines for Therapeutic or Prophylactic Purposes

Pharmaceutical and vaccine compositions of the invention are typically used to treat and/or prevent a cancer that expresses or overexpresses 158P1D7. In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective B cell, CTL and/or HTL response to the antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

For pharmaceutical compositions, the immunogenic peptides of the invention, or DNA encoding them, are generally administered to an individual already bearing a tumor that expresses 158P1D7. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences. Patients can be treated with the immunogenic peptides separately or in conjunction with other treatments, such as surgery, as appropriate.

For therapeutic use, administration should generally begin at the first diagnosis of 158P1D7-associated cancer. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. The embodiment of the vaccine composition (*i.e.*, including, but not limited to embodiments such as peptide cocktails, polyepitopic polypeptides, minigenes, or TAA-specific CTLs or pulsed dendritic cells) delivered to the patient may vary according to the stage of the disease or the patient's health status. For example, in a patient with a tumor that expresses 158P1D7, a vaccine comprising 158P1D7-specific CTL may be more efficacious in killing tumor cells in patient with advanced disease than alternative embodiments.

It is generally important to provide an amount of the peptide epitope delivered by a mode of administration sufficient to effectively stimulate a cytotoxic T cell response; compositions which stimulate helper T cell responses can also be given in accordance with this embodiment of the invention.

The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1,000 µg and the higher value is about 10,000; 20,000; 30,000; or 50,000 µg. Dosage values for a human typically range from about 500 µg to about 50,000 µg per 70 kilogram patient. Boosting dosages of between about 1.0 µg to about 50,000 µg of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and HTL obtained from the patient's blood. Administration should continue until at least clinical symptoms or laboratory tests indicate that the

neoplasia, has been eliminated or reduced and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

In certain embodiments, the peptides and compositions of the present invention are employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

The vaccine compositions of the invention can also be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 ßg and the higher value is about 10,000; 20,000; 30,000; or 50,000 ßg. Dosage values for a human typically range from about 500 ßg to about 50,000 ßg per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 µg to about 50,000 µg of peptide administered at defined intervals from about four weeks to six months after the initial administration of vaccine. The immunogenicity of the vaccine can be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, nasal, intrathecal, or local (*e.g.* as a cream or topical ointment) administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously, subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier.

A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well-known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration.

The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g.*, Remington's Pharmaceutical Sciences, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

Proteins(s) of the invention, and/or nucleic acids encoding the protein(s), can also be administered via liposomes, which may also serve to: 1) target the proteins(s) to a particular tissue, such as lymphoid tissue; 2) to target selectively to diseases cells; or, 3) to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention

can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

For aerosol administration, immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are about 0.01%-20% by weight, preferably about 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from about 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute about 0.1%-20% by weight of the composition, preferably about 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, *e.g.*, lecithin for intranasal delivery.

XI.) Diagnostic and Prognostic Embodiments of 158P1D7.

As disclosed herein, 158P1D7 polynucleotides, polypeptides, reactive cytotoxic T cells (CTL), reactive helper T cells (HTL) and anti-polypeptide antibodies are used in well known diagnostic, prognostic and therapeutic assays that examine conditions associated with dysregulated cell growth such as cancer, in particular the cancers listed in Table I (see, e.g., both its specific pattern of tissue expression as well as its overexpression in certain cancers as described for example in Example 4).

158P1D7 can be used in a manner analogous to, or as complementary to, the bladder associated antigen combination, mucins and CEA, represented in a diagnostic kit called ImmunoCyt™. ImmunoCyt a is a commercially available assay to identify and monitor the presence of bladder cancer (see Fradet et al., 1997, Can J Urol, 4(3):400-405). A variety of other diagnostic markers are also used in similar contexts including p53 and H-ras (see, e.g., Tulchinsky et al., Int J Mol Med 1999 Jul 4(1):99-102 and Minimoto et al., Cancer Detect Prev 2000;24(1):1-12). Therefore, this disclosure of the 158P1D7 polynucleotides and polypeptides (as well as the 158P1D7 polynucleotide probes and anti-158P1D7 antibodies used to identify the presence of these molecules) and their properties allows skilled artisans to utilize these molecules in methods that are analogous to those used, for example, in a variety of diagnostic assays directed to examining conditions associated with cancer.

Typical embodiments of diagnostic methods which utilize the 158P1D7 polynucleotides, polypeptides, reactive T cells and antibodies are analogous to those methods from well-established diagnostic assays which employ, e.g., PSA polynucleotides, polypeptides, reactive T cells and antibodies. For example, just as PSA polynucleotides are used as probes

(for example in Northern analysis, see, e.g., Sharief et al., Biochem. Mol. Biol. Int. 33(3):567-74(1994)) and primers (for example in PCR analysis, see, e.g., Okegawa et al., J. Urol. 163(4): 1189-1190 (2000)) to observe the presence and/or the level of PSA mRNAs in methods of monitoring PSA overexpression or the metastasis of prostate cancers, the 158P1D7 polynucleotides described herein can be utilized to detect 158P1D7 overexpression or the metastasis of bladder and other cancers expressing this gene. Alternatively, just as PSA polypeptides are used to generate antibodies specific for PSA which can then be used to observe the presence and/or the level of PSA proteins in methods to monitor PSA protein overexpression (see, e.g., Stephan et al., Urology 55(4):560-3 (2000)) or the metastasis of prostate cells (see, e.g., Alanen et al., Pathol. Res. Pract. 192(3):233-7 (1996)), the 158P1D7 polypeptides described herein can be utilized to generate antibodies for use in detecting 158P1D7 overexpression or the metastasis of bladder cells and cells of other cancers expressing this gene.

Specifically, because metastases involves the movement of cancer cells from an organ of origin (such as the lung or bladder etc.) to a different area of the body (such as a lymph node), assays which examine a biological sample for the presence of cells expressing 158P1D7 polynucleotides and/or polypeptides can be used to provide evidence of metastasis. For example, when a biological sample from tissue that does not normally contain 158P1D7-expressing cells (lymph node) is found to contain 158P1D7-expressing cells such as the 158P1D7 expression seen in LAPC4 and LAPC9, xenografts isolated from lymph node and bone metastasis, respectively, this finding is indicative of metastasis.

Alternatively 158P1D7 polynucleotides and/or polypeptides can be used to provide evidence of cancer, for example, when cells in a biological sample that do not normally express 158P1D7 or express 158P1D7 at a different level are found to express 158P1D7 or have an increased expression of 158P1D7 (see, e.g., the 158P1D7 expression in the cancers listed in Table I and in patient samples etc. shown in the accompanying Figures). In such assays, artisans may further wish to generate supplementary evidence of metastasis by testing the biological sample for the presence of a second tissue restricted marker (in addition to 158P1D7) such as ImmunoCyt™, PSCA etc. (see, e.g., Fradet et al., 1997, Can J Urol, 4(3):400-405; Amara et al., 2001, Cancer Res 61:4660-4665). Just as PSA polynucleotide fragments and polynucleotide variants are employed by skilled artisans for use in methods of monitoring PSA, 158P1D7 polynucleotide fragments and polynucleotide variants are used in an analogous manner. In particular, typical PSA polynucleotides used in methods of monitoring PSA are probes or primers which consist of fragments of the PSA cDNA sequence. Illustrating this, primers used to PCR amplify a PSA polynucleotide must include less than the whole PSA sequence to function in the polymerase chain reaction. In the context of such PCR reactions, skilled artisans generally create a variety of different polynucleotide fragments that can be used as primers in order to amplify different portions of a polynucleotide of interest or to optimize amplification reactions (see, e.g., Caetano-Anolles, G. Biotechniques 25(3): 472-476, 478-480 (1998); Robertson et al., Methods Mol. Biol. 98:121-154 (1998)). An additional illustration of the use of such fragments is provided in Example 4, where a 158P1D7 polynucleotide fragment is used as a probe to show the expression of 158P1D7 RNAs in cancer cells. In addition, variant polynucleotide sequences are typically used as primers and probes for the corresponding mRNAs in PCR and Northern analyses (see, e.g., Sawai et al., Fetal Diagn. Ther. 1996 Nov-Dec 11(6):407-13 and Current Protocols In Molecular Biology, Volume 2, Unit 2, Frederick M. Ausubel et al. eds., 1995)). Polynucleotide fragments and variants are useful in this context where they are capable of binding to a target polynucleotide sequence (e.g. the 158P1D7 polynucleotide shown in Figure 2) under conditions of high stringency.

Furthermore, PSA polypeptides which contain an epitope that can be recognized by an antibody or T cell that specifically binds to that epitope are used in methods of monitoring PSA. 158P1D7 polypeptide fragments and polypeptide analogs or variants can also be used in an analogous manner. This practice of using polypeptide fragments or polypeptide variants to generate antibodies (such as anti-PSA antibodies or T cells) is typical in the art with a wide variety of systems such as fusion proteins being used by practitioners (see, e.g., Current Protocols In Molecular Biology, Volume 2, Unit 16,

Frederick M. Ausubel et al. eds., 1995). In this context, each epitope(s) functions to provide the architecture with which an antibody or T cell is reactive. Typically, skilled artisans create a variety of different polypeptide fragments that can be used in order to generate immune responses specific for different portions of a polypeptide of interest (see, e.g., U.S. Patent No. 5,840,501 and U.S. Patent No. 5,939,533). For example it may be preferable to utilize a polypeptide comprising one of the 158P1D7 biological motifs discussed herein or a motif-bearing subsequence which is readily identified by one of skill in the art based on motifs available in the art. Polypeptide fragments, variants or analogs are typically useful in this context as long as they comprise an epitope capable of generating an antibody or T cell specific for a target polypeptide sequence (e.g. the 158P1D7 polypeptide shown in Figure 2).

As shown herein, the 158P1D7 polynucleotides and polypeptides (as well as the 158P1D7 polynucleotide probes and anti-158P1D7 antibodies or T cells used to identify the presence of these molecules) exhibit specific properties that make them useful in diagnosing cancers such as those listed in Table I. Diagnostic assays that measure the presence of 158P1D7 gene products, in order to evaluate the presence or onset of a disease condition described herein, such as bladder cancer, are used to identify patients for preventive measures or further monitoring, as has been done so successfully with PSA for monitoring prostate cancer. Materials such as 158P1D7 polynucleotides and polypeptides (as well as the 158P1D7 polynucleotide probes and anti-158P1D7 antibodies used to identify the presence of these molecules) satisfy a need in the art for molecules having similar or complementary characteristics to PSA in situations of bladder cancer. Finally, in addition to their use in diagnostic assays, the 158P1D7 polynucleotides disclosed herein have a number of other utilities such as their use in the identification of oncogenetic associated chromosomal abnormalities in the chromosomal region to which the 158P1D7 gene maps (see Example 3 below). Moreover, in addition to their use in diagnostic assays, the 158P1D7-related proteins and polynucleotides disclosed herein have other utilities such as their use in the forensic analysis of tissues of unknown origin (see, e.g., Takahama K Forensic Sci Int 1996 Jun 28;80(1-2): 63-9).

Additionally, 158P1D7-related proteins or polynucleotides of the invention can be used to treat a pathologic condition characterized by the over-expression of 158P1D7. For example, the amino acid or nucleic acid sequence of Figure 2 or Figure 3, or fragments of either, can be used to generate an immune response to the 158P1D7 antigen. Antibodies or other molecules that react with 158P1D7 can be used to modulate the function of this molecule, and thereby provide a therapeutic benefit.

XII.) Inhibition of 158P1D7 Protein Function

The invention includes various methods and compositions for inhibiting the binding of 158P1D7 to its binding partner or its association with other protein(s) as well as methods for inhibiting 158P1D7 function.

XII.A.) Inhibition of 158P1D7 With Intracellular Antibodies

In one approach, a recombinant vector that encodes single chain antibodies that specifically bind to 158P1D7 are introduced into 158P1D7 expressing cells via gene transfer technologies. Accordingly, the encoded single chain anti-158P1D7 antibody is expressed intracellularly, binds to 158P1D7 protein, and thereby inhibits its function. Methods for engineering such intracellular single chain antibodies are well known. Such intracellular antibodies, also known as "intrabodies", are specifically targeted to a particular compartment within the cell, providing control over where the inhibitory activity of the treatment is focused. This technology has been successfully applied in the art (for review, see Richardson and Marasco, 1995, TIBTECH vol. 13). Intrabodies have been shown to virtually eliminate the expression of otherwise abundant cell surface receptors (see, e.g., Richardson et al., 1995, Proc. Natl. Acad. Sci. USA 92: 3137-3141; Beerli et al., 1994, J. Biol. Chem. 289: 23931-23936; Deshane et al., 1994, Gene Ther. 1: 332-337).

Single chain antibodies comprise the variable domains of the heavy and light chain joined by a flexible linker polypeptide, and are expressed as a single polypeptide. Optionally, single chain antibodies are expressed as a single chain

variable region fragment joined to the light chain constant region. Well-known intracellular trafficking signals are engineered into recombinant polynucleotide vectors encoding such single chain antibodies in order to precisely target the intrabody to the desired intracellular compartment. For example, intrabodies targeted to the endoplasmic reticulum (ER) are engineered to incorporate a leader peptide and, optionally, a C-terminal ER retention signal, such as the KDEL amino acid motif. Intrabodies intended to exert activity in the nucleus are engineered to include a nuclear localization signal. Lipid moieties are joined to intrabodies in order to tether the intrabody to the cytosolic side of the plasma membrane. Intrabodies can also be targeted to exert function in the cytosol. For example, cytosolic intrabodies are used to sequester factors within the cytosol, thereby preventing them from being transported to their natural cellular destination.

In one embodiment, intrabodies are used to capture 158P1D7 in the nucleus, thereby preventing its activity within the nucleus. Nuclear targeting signals are engineered into such 158P1D7 intrabodies in order to achieve the desired targeting. Such 158P1D7 intrabodies are designed to bind specifically to a particular 158P1D7 domain. In another embodiment, cytosolic intrabodies that specifically bind to the 158P1D7 protein are used to prevent 158P1D7 from gaining access to the nucleus, thereby preventing it from exerting any biological activity within the nucleus (e.g., preventing 158P1D7 from forming transcription complexes with other factors).

In order to specifically direct the expression of such intrabodies to particular cells, the transcription of the intrabody is placed under the regulatory control of an appropriate tumor-specific promoter and/or enhancer. In order to target intrabody expression specifically to bladder, for example, the PSCA promoter and/or promoter/enhancer can be utilized (See, for example, U.S. Patent No. 5,919,652 issued 6 July 1999 and Lin et al. PNAS, USA 92(3):679-683 (1995)).

XII.B.) Inhibition of 158P1D7 with Recombinant Proteins

In another approach, recombinant molecules bind to 158P1D7 and thereby inhibit 158P1D7 function. For example, these recombinant molecules prevent or inhibit 158P1D7 from accessing/binding to its binding partner(s) or associating with other protein(s). Such recombinant molecules can, for example, contain the reactive part(s) of a 158P1D7 specific antibody molecule. In a particular embodiment, the 158P1D7 binding domain of a 158P1D7 binding partner is engineered into a dimeric fusion protein, whereby the fusion protein comprises two 158P1D7 ligand binding domains linked to the Fc portion of a human IgG, such as human IgG1. Such IgG portion can contain, for example, the C_H2 and C_H3 domains and the hinge region, but not the C_H1 domain. Such dimeric fusion proteins are administered in soluble form to patients suffering from a cancer associated with the expression of 158P1D7, whereby the dimeric fusion protein specifically binds to 158P1D7 and blocks 158P1D7 interaction with a binding partner. Such dimeric fusion proteins are further combined into multimeric proteins using known antibody linking technologies.

XII.C.) Inhibition of 158P1D7 Transcription or Translation

The present invention also comprises various methods and compositions for inhibiting the transcription of the 158P1D7 gene. Similarly, the invention also provides methods and compositions for inhibiting the translation of 158P1D7 mRNA into protein.

In one approach, a method of inhibiting the transcription of the 158P1D7 gene comprises contacting the 158P1D7 gene with a 158P1D7 antisense polynucleotide. In another approach, a method of inhibiting 158P1D7 mRNA translation comprises contacting the 158P1D7 mRNA with an antisense polynucleotide. In another approach, a 158P1D7 specific ribozyme is used to cleave the 158P1D7 message, thereby inhibiting translation. Such antisense and ribozyme based methods can also be directed to the regulatory regions of the 158P1D7 gene, such as the 158P1D7 promoter and/or enhancer elements. Similarly, proteins capable of inhibiting a 158P1D7 gene transcription factor are used to inhibit 158P1D7 mRNA transcription. The various polynucleotides and compositions useful in the aforementioned methods have been described above. The use of antisense and ribozyme molecules to inhibit transcription and translation is well known in the art.

Other factors that inhibit the transcription of 158P1D7 by interfering with 158P1D7 transcriptional activation are also useful to treat cancers expressing 158P1D7. Similarly, factors that interfere with 158P1D7 processing are useful to treat cancers that express 158P1D7. Cancer treatment methods utilizing such factors are also within the scope of the invention.

XII.D.) General Considerations for Therapeutic Strategies

Gene transfer and gene therapy technologies can be used to deliver therapeutic polynucleotide molecules to tumor cells synthesizing 158P1D7 (i.e., antisense, ribozyme, polynucleotides encoding intrabodies and other 158P1D7 inhibitory molecules). A number of gene therapy approaches are known in the art. Recombinant vectors encoding 158P1D7 antisense polynucleotides, ribozymes, factors capable of interfering with 158P1D7 transcription, and so forth, can be delivered to target tumor cells using such gene therapy approaches.

The above therapeutic approaches can be combined with any one of a wide variety of surgical, chemotherapy or radiation therapy regimens. The therapeutic approaches of the invention can enable the use of reduced dosages of chemotherapy (or other therapies) and/or less frequent administration, an advantage for all patients and particularly for those that do not tolerate the toxicity of the chemotherapeutic agent well.

The anti-tumor activity of a particular composition (e.g., antisense, ribozyme, intrabody), or a combination of such compositions, can be evaluated using various *in vitro* and *in vivo* assay systems. *In vitro* assays that evaluate therapeutic activity include cell growth assays, soft agar assays and other assays indicative of tumor promoting activity, binding assays capable of determining the extent to which a therapeutic composition will inhibit the binding of 158P1D7 to a binding partner, etc.

In vivo, the effect of a 158P1D7 therapeutic composition can be evaluated in a suitable animal model. For example, xenogenic bladder cancer models can be used, wherein human bladder cancer explants or passaged xenograft tissues are introduced into immune compromised animals, such as nude or SCID mice (Shibayama et al., 1991, J Urol., 146(4):1136-7; Beecken et al., 2000, Urology, 56(3):521-526). Efficacy can be predicted using assays that measure inhibition of tumor formation, tumor regression or metastasis, and the like.

In vivo assays that evaluate the promotion of apoptosis are useful in evaluating therapeutic compositions. In one embodiment, xenografts from tumor bearing mice treated with the therapeutic composition can be examined for the presence of apoptotic foci and compared to untreated control xenograft-bearing mice. The extent to which apoptotic foci are found in the tumors of the treated mice provides an indication of the therapeutic efficacy of the composition.

The therapeutic compositions used in the practice of the foregoing methods can be formulated into pharmaceutical compositions comprising a carrier suitable for the desired delivery method. Suitable carriers include any material that when combined with the therapeutic composition retains the anti-tumor function of the therapeutic composition and is generally non-reactive with the patient's immune system. Examples include, but are not limited to, any of a number of standard pharmaceutical carriers such as sterile phosphate buffered saline solutions, bacteriostatic water, and the like (see, generally, Remington's Pharmaceutical Sciences 16th Edition, A. Osal., Ed., 1980).

Therapeutic formulations can be solubilized and administered via any route capable of delivering the therapeutic composition to the tumor site. Potentially effective routes of administration include, but are not limited to, intravenous, parenteral, intraperitoneal, intramuscular, intratumor, intradermal, intraorgan, orthotopic, and the like. A preferred formulation for intravenous injection comprises the therapeutic composition in a solution of preserved bacteriostatic water, sterile unpreserved water, and/or diluted in polyvinylchloride or polyethylene bags containing 0.9% sterile Sodium Chloride for Injection, USP. Therapeutic protein preparations can be lyophilized and stored as sterile powders, preferably under vacuum, and then reconstituted in bacteriostatic water (containing for example, benzyl alcohol preservative) or in sterile water prior to injection.

Dosages and administration protocols for the treatment of cancers using the foregoing methods will vary with the method and the target cancer, and will generally depend on a number of other factors appreciated in the art.

XIII.) Identification, Characterization and Use of Modulators of 158P1D7

Methods to Identify and Use Modulators

In one embodiment, screening is performed to identify modulators that induce or suppress a particular expression profile, suppress or induce specific pathways, preferably generating the associated phenotype thereby. In another embodiment, having identified differentially expressed genes important in a particular state; screens are performed to identify modulators that alter expression of individual genes, either increase or decrease. In another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition, screens are done for genes that are induced in response to a candidate agent. After identifying a modulator (one that suppresses a cancer expression pattern leading to a normal expression pattern, or a modulator of a cancer gene that leads to expression of the gene as in normal tissue) a screen is performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent-treated cancer tissue reveals genes that are not expressed in normal tissue or cancer tissue, but are expressed in agent treated tissue, and vice versa. These agent-specific sequences are identified and used by methods described herein for cancer genes or proteins. In particular these sequences and the proteins they encode are used in marking or identifying agent-treated cells. In addition, antibodies are raised against the agent-induced proteins and used to target novel therapeutics to the treated cancer tissue sample.

Modulator-related Identification and Screening Assays:

Gene Expression-related Assays

Proteins, nucleic acids, and antibodies of the invention are used in screening assays. The cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing these sequences are used in screening assays, such as evaluating the effect of drug candidates on a "gene expression profile," expression profile of polypeptides or alteration of biological function. In one embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Davis, GF, et al, J Biol Screen 7:69 (2002); Zlokarnik, et al., Science 279:84-8 (1998); Heid, Genome Res 6:986-94,1996).

The cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified cancer proteins or genes are used in screening assays. That is, the present invention comprises methods for screening for compositions which modulate the cancer phenotype or a physiological function of a cancer protein of the invention. This is done on a gene itself or by evaluating the effect of drug candidates on a "gene expression profile" or biological function. In one embodiment, expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring after treatment with a candidate agent, see Zlokarnik, supra.

A variety of assays are executed directed to the genes and proteins of the invention. Assays are run on an individual nucleic acid or protein level. That is, having identified a particular gene as up regulated in cancer, test compounds are screened for the ability to modulate gene expression or for binding to the cancer protein of the invention. "Modulation" in this context includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold

decrease in cancer tissue compared to normal tissue a target value of a 10-fold increase in expression by the test compound is often desired. Modulators that exacerbate the type of gene expression seen in cancer are also useful, e.g., as an upregulated target in further analyses.

The amount of gene expression is monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, a gene product itself is monitored, e.g., through the use of antibodies to the cancer protein and standard immunoassays. Proteomics and separation techniques also allow for quantification of expression.

Expression Monitoring to Identify Compounds that Modify Gene Expression

In one embodiment, gene expression monitoring, i.e., an expression profile, is monitored simultaneously for a number of entities. Such profiles will typically involve one or more of the genes of Figure 2. In this embodiment, e.g., cancer nucleic acid probes are attached to biochips to detect and quantify cancer sequences in a particular cell. Alternatively, PCR can be used. Thus, a series, e.g., wells of a microtiter plate, can be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring is performed to identify compounds that modify the expression of one or more cancer-associated sequences, e.g., a polynucleotide sequence set out in Figure 2. Generally, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate cancer, modulate cancer proteins of the invention, bind to a cancer protein of the invention, or interfere with the binding of a cancer protein of the invention and an antibody or other binding partner.

In one embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds," as compounds for screening, or as therapeutics.

In certain embodiments, combinatorial libraries of potential modulators are screened for an ability to bind to a cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

As noted above, gene expression monitoring is conveniently used to test candidate modulators (e.g., protein, nucleic acid or small molecule). After the candidate agent has been added and the cells allowed to incubate for a period, the sample containing a target sequence to be analyzed is, e.g., added to a biochip.

If required, the target sequence is prepared using known techniques. For example, a sample is treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an in vitro transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

The target sequence can be labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that is detected. Alternatively, the label is a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5, 681,702; 5,597,909; 5,545,730; 5,594,117; 5,591,584; 5,571,670; 5,580,731; 5,571,670; 5,591,584; 5,624,802; 5,635,352; 5,594,118; 5,359,100; 5,124, 246; and 5,681,697. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions are used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc. These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus, it can be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein can be accomplished in a variety of ways. Components of the reaction can be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g. albumin, detergents, etc. which can be used to facilitate optimal hybridization and detection, and/or reduce nonspecific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target. The assay data are analyzed to determine the expression levels of individual genes, and changes in expression levels as between states, forming a gene expression profile.

Biological Activity-related Assays

The invention provides methods identify or screen for a compound that modulates the activity of a cancer-related gene or protein of the invention. The methods comprise adding a test compound, as defined above, to a cell comprising a cancer protein of the invention. The cells contain a recombinant nucleic acid that encodes a cancer protein of the invention. In another embodiment, a library of candidate agents is tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g. hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e., cell-cell contacts). In another example, the determinations are made at different stages of the cell cycle process. In this way, compounds that modulate genes or proteins of the invention are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the cancer protein of the invention. Once identified, similar structures are evaluated to identify critical structural features of the compound.

In one embodiment, a method of modulating (e.g., inhibiting) cancer cell division is provided; the method comprises administration of a cancer modulator. In another embodiment, a method of modulating (e.g., inhibiting) cancer is provided; the method comprises administration of a cancer modulator. In a further embodiment, methods of treating cells or individuals with cancer are provided; the method comprises administration of a cancer modulator.

In one embodiment, a method for modulating the status of a cell that expresses a gene of the invention is provided. As used herein status comprises such art-accepted parameters such as growth, proliferation, survival, function, apoptosis, senescence, location, enzymatic activity, signal transduction, etc. of a cell. In one embodiment, a cancer inhibitor is an

antibody as discussed above. In another embodiment, the cancer inhibitor is an antisense molecule. A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described herein.

High Throughput Screening to Identify Modulators

The assays to identify suitable modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

In one embodiment, modulators evaluated in high throughput screening methods are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, are used. In this way, libraries of proteins are made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes, or ligands and receptors.

Use of Soft Agar Growth and Colony Formation to Identify and Characterize Modulators

Normal cells require a solid substrate to attach and grow. When cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, can regenerate normal phenotype and once again require a solid substrate to attach to and grow. Soft agar growth or colony formation in assays are used to identify modulators of cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A modulator reduces or eliminates the host cells' ability to grow suspended in solid or semisolid media, such as agar.

Techniques for soft agar growth or colony formation in suspension assays are described in Freshney, Culture of Animal Cells a Manual of Basic Technique (3rd ed., 1994). See also, the methods section of Garkavtsev et al. (1996), supra.

Evaluation of Contact Inhibition and Growth Density Limitation to Identify and Characterize Modulators

Normal cells typically grow in a flat and organized pattern in cell culture until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. Transformed cells, however, are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, transformed cells grow to a higher saturation density than corresponding normal cells. This is detected morphologically by the formation of a disoriented monolayer of cells or cells in foci. Alternatively, labeling index with (^{3}H)-thymidine at saturation density is used to measure density limitation of growth, similarly an MTT or Alamar blue assay will reveal proliferation capacity of cells and the the ability of modulators to affect same. See Freshney (1994), supra. Transformed cells, when transfected with tumor suppressor genes, can regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with ^{3}H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (^{3}H)-thymidine is determined by incorporated cpm.

Contact independent growth is used to identify modulators of cancer sequences, which had led to abnormal cellular proliferation and transformation. A modulator reduces or eliminates contact independent growth, and returns the cells to a normal phenotype.

Evaluation of Growth Factor or Serum Dependence to Identify and Characterize Modulators

Transformed cells have lower serum dependence than their normal counterparts (see, e.g., Temin, J. Natl. Cancer Inst. 37:167-175 (1966); Eagle et al., J. Exp. Med 131:836-879 (1970)); Freshney, supra. This is in part due to release of

various growth factors by the transformed cells. The degree of growth factor or serum dependence of transformed host cells can be compared with that of control. For example, growth factor or serum dependence of a cell is monitored in methods to identify and characterize compounds that modulate cancer-associated sequences of the invention.

Use of Tumor-specific Marker Levels to Identify and Characterize Modulators

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g., Gullino, Angiogenesis, Tumor Vascularization, and Potential Interference with Tumor Growth, *in* Biological Responses in Cancer, pp. 178-184 (Mihich (ed.) 1985)). Similarly, Tumor Angiogenesis Factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman, Angiogenesis and Cancer, Sem Cancer Biol. (1992)), while bFGF is released from endothelial tumors (Ensoli, B et al).

Various techniques which measure the release of these factors are described in Freshney (1994), supra. Also, see, Unkless et al., J. Biol. Chem. 249:4295-4305 (1974); Strickland & Beers, J. Biol. Chem. 251:5694-5702 (1976); Whur et al., Br. J. Cancer 42:305 312 (1980); Gullino, Angiogenesis, Tumor Vascularization, and Potential Interference with Tumor Growth, *in* Biological Responses in Cancer, pp. 178-184 (Mihich (ed.) 1985); Freshney, Anticancer Res. 5:111-130 (1985). For example, tumor specific marker levels are monitored in methods to identify and characterize compounds that modulate cancer-associated sequences of the invention.

Invasiveness into Matrigel to Identify and Characterize Modulators

The degree of invasiveness into Matrigel or an extracellular matrix constituent can be used as an assay to identify and characterize compounds that modulate cancer associated sequences. Tumor cells exhibit a positive correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells. Techniques described in Cancer Res. 1999; 59:6010; Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells is measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{1251} and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), supra.

Evaluation of Tumor Growth *In Vivo* to Identify and Characterize Modulators

Effects of cancer-associated sequences on cell growth are tested in transgenic or immune-suppressed organisms. Transgenic organisms are prepared in a variety of art-accepted ways. For example, knock-out transgenic organisms, e.g., mammals such as mice, are made, in which a cancer gene is disrupted or in which a cancer gene is inserted. Knock-out transgenic mice are made by insertion of a marker gene or other heterologous gene into the endogenous cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous cancer gene with a mutated version of the cancer gene, or by mutating the endogenous cancer gene, e.g., by exposure to carcinogens.

To prepare transgenic chimeric animals, e.g., mice, a DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells some of which are derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi et al., Science 244:1288 (1989)). Chimeric mice can be derived according to US Patent 6,365,797, issued 2 April 2002; US Patent 6,107,540 issued 22 August 2000; Hogan et al., Manipulating the Mouse Embryo: A laboratory Manual, Cold Spring Harbor Laboratory (1988) and Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, ed., IRL Press, Washington, D.C., (1987).

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, a genetically athymic "nude" mouse (see, e.g., Giovanella et al., J. Natl. Cancer Inst. 52:921 (1974)), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley et al., Br. J. Cancer 38:263 (1978); Selby et al., Br. J. Cancer 41:52 (1980)) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts produce invasive tumors in a high proportion of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing cancer-associated sequences are injected subcutaneously or orthotopically. Mice are then separated into groups, including control groups and treated experimental groups) e.g. treated with a modulator). After a suitable length of time, preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions, or weight) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

In Vitro Assays to Identify and Characterize Modulators

Assays to identify compounds with modulating activity can be performed in vitro. For example, a cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the cancer polypeptide levels are determined in vitro by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as Western blotting, ELISA and the like with an antibody that selectively binds to the cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e. g., Northern hybridization, RNAse protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or P-gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art (Davis GF, supra; Gonzalez, J. & Negulescu, P. Curr. Opin. Biotechnol. 1998: 9:624).

As outlined above, in vitro screens are done on individual genes and gene products. That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself is performed.

In one embodiment, screening for modulators of expression of specific gene(s) is performed. Typically, the expression of only one or a few genes is evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

Binding Assays to Identify and Characterize Modulators

In binding assays in accordance with the invention, a purified or isolated gene product of the invention is generally used. For example, antibodies are generated to a protein of the invention, and immunoassays are run to determine the amount and/or location of protein. Alternatively, cells comprising the cancer proteins are used in the assays.

Thus, the methods comprise combining a cancer protein of the invention and a candidate compound such as a ligand, and determining the binding of the compound to the cancer protein of the invention. Preferred embodiments utilize the human cancer protein; animal models of human disease of can also be developed and used. Also, other analogous

mammalian proteins also can be used as appreciated by those of skill in the art. Moreover, in some embodiments variant or derivative cancer proteins are used.

Generally, the cancer protein of the invention, or the ligand, is non-diffusibly bound to an insoluble support. The support can, e.g., be one having isolated sample receiving areas (a microtiter plate, an array, etc.). The insoluble supports can be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports can be solid or porous and of any convenient shape.

Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharide, nylon, nitrocellulose, or Teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition to the support is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies which do not sterically block either the ligand binding site or activation sequence when attaching the protein to the support, direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or ligand/binding agent to the support, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

Once a cancer protein of the invention is bound to the support, and a test compound is added to the assay. Alternatively, the candidate binding agent is bound to the support and the cancer protein of the invention is then added. Binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc.

Of particular interest are assays to identify agents that have a low toxicity for human cells. A wide variety of assays can be used for this purpose, including proliferation assays, cAMP assays, labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

A determination of binding of the test compound (ligand, binding agent, modulator, etc.) to a cancer protein of the invention can be done in a number of ways. The test compound can be labeled, and binding determined directly, e.g., by attaching all or a portion of the cancer protein of the invention to a solid support, adding a labeled candidate compound (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps can be utilized as appropriate.

In certain embodiments, only one of the components is labeled, e.g., a protein of the invention or ligands labeled. Alternatively, more than one component is labeled with different labels, e.g., I^{125}, for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

Competitive Binding to Identify and Characterize Modulators

In one embodiment, the binding of the "test compound" is determined by competitive binding assay with a "competitor." The competitor is a binding moiety that binds to the target molecule (e.g., a cancer protein of the invention). Competitors include compounds such as antibodies, peptides, binding partners, ligands, etc. Under certain circumstances, the competitive binding between the test compound and the competitor displaces the test compound. In one embodiment, the test compound is labeled. Either the test compound, the competitor, or both, is added to the protein for a time sufficient to allow binding. Incubations are performed at a temperature that facilitates optimal activity, typically between four and 40°C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening; typically between zero and one

hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In one embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the cancer protein and thus is capable of binding to, and potentially modulating, the activity of the cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the post-test compound wash solution indicates displacement by the test compound. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor indicates that the test compound binds to the cancer protein with higher affinity than the competitor. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, indicates that the test compound binds to and thus potentially modulates the cancer protein of the invention.

Accordingly, the competitive binding methods comprise differential screening to identity agents that are capable of modulating the activity of the cancer proteins of the invention. In this embodiment, the methods comprise combining a cancer protein and a competitor in a first sample. A second sample comprises a test compound, the cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native cancer protein, but cannot bind to modified cancer proteins. For example the structure of the cancer protein is modeled and used in rational drug design to synthesize agents that interact with that site, agents which generally do not bind to site-modified proteins. Moreover, such drug candidates that affect the activity of a native cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of such proteins.

Positive controls and negative controls can be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples occurs for a time sufficient to allow for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples can be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents can be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. which are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., can be used. The mixture of components is added in an order that provides for the requisite binding.

Use of Polynucleotides to Down-regulate or Inhibit a Protein of the Invention.

Polynucleotide modulators of cancer can be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand-binding molecule, as described in WO 91/04753. Suitable ligand-binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of cancer can be introduced into a cell

containing the target nucleic acid sequence, e.g., by formation of a polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Inhibitory and Antisense Nucleotides

In certain embodiments, the activity of a cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide or inhibitory small nuclear RNA (snRNA), i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a cancer protein of the invention, mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally occurring nucleotides, or synthetic species formed from naturally occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprised by this invention so long as they function effectively to hybridize with nucleotides of the invention. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized in vitro. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 12 nucleotides, preferably from about 12 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein &Cohen (Cancer Res. 48:2659 (1988 and van der Krol et al. (BioTechniques 6:958 (1988)).

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto et al., Adv. in Pharmacology 25: 289-317 (1994) for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel et al., Nucl. Acids Res. 18:299-304 (1990); European Patent Publication No. 0360257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang et al., Proc. Natl. Acad. Sci. USA 90:6340-6344 (1993); Yamada et al., Human Gene Therapy 1:39-45 (1994); Leavitt et al., Proc. Natl. Acad Sci. USA 92:699- 703 (1995); Leavitt et al., Human Gene Therapy 5: 1151-120 (1994); and Yamada et al., Virology 205: 121-126 (1994)).

Use of Modulators in Phenotypic Screening

In one embodiment, a test compound is administered to a population of cancer cells, which have an associated cancer expression profile. By "administration" or "contacting" herein is meant that the modulator is added to the cells in such a manner as to allow the modulator to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, a nucleic acid encoding a proteinaceous agent (i.e., a peptide) is put into a viral construct

such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used. Once the modulator has been administered to the cells, the cells are washed if desired and are allowed to incubate under preferably physiological conditions for some period. The cells are then harvested and a new gene expression profile is generated. Thus, e.g., cancer tissue is screened for agents that modulate, e.g., induce or suppress, the cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on cancer activity. Similarly, altering a biological function or a signaling pathway is indicative of modulator activity. By defining such a signature for the cancer phenotype, screens for new drugs that alter the phenotype are devised. With this approach, the drug target need not be known and need not be represented in the original gene/protein expression screening platform, nor does the level of transcript for the target protein need to change. The modulator inhibiting function will serve as a surrogate marker

As outlined above, screens are done to assess genes or gene products. That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself is performed.

Use of Modulators to Affect Peptides of the Invention

Measurements of cancer polypeptide activity, or of the cancer phenotype are performed using a variety of assays. For example, the effects of modulators upon the function of a cancer polypeptide(s) are measured by examining parameters described above. A physiological change that affects activity is used to assess the influence of a test compound on the polypeptides of this invention. When the functional outcomes are determined using intact cells or animals, a variety of effects can be assesses such as, in the case of a cancer associated with solid tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., by Northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGNIP.

Methods of Identifying Characterizing Cancer-associated Sequences

Expression of various gene sequences is correlated with cancer. Accordingly, disorders based on mutant or variant cancer genes are determined. In one embodiment, the invention provides methods for identifying cells containing variant cancer genes, e.g., determining the presence of, all or part, the sequence of at least one endogenous cancer gene in a cell. This is accomplished using any number of sequencing techniques. The invention comprises methods of identifying the cancer genotype of an individual, e.g., determining all or part of the sequence of at least one gene of the invention in the individual. This is generally done in at least one tissue of the individual, e.g., a tissue set forth in Table I, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced gene to a known cancer gene, i.e., a wild-type gene to determine the presence of family members, homologies, mutations or variants. The sequence of all or part of the gene can then be compared to the sequence of a known cancer gene to determine if any differences exist. This is done using any number of known homology programs, such as BLAST, Bestfit, etc. The presence of a difference in the sequence between the cancer gene of the patient and the known cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the cancer genes are used as probes to determine the number of copies of the cancer gene in the genome. The cancer genes are used as probes to determine the chromosomal localization of the cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the cancer gene locus.

XIV.) RNAi and Therapeutic Use of Small Interfering RNA (siRNAs)

The present invention is also directed towards siRNA oligonucleotides, particularly double stranded RNAs encompassing at least a fragment of the 158P1D7 coding region or 5" UTR regions, or complement, or any antisense oligonucleotide specific to the 158P1D7 sequence. In one embodiment such oligonucleotides are used to elucidate a function of 158P1D7, or are used to screen for or evaluate modulators of 158P1D7 function or expression. In another embodiment, gene expression of 158P1D7 is reduced by using siRNA transfection and results in significantly diminished proliferative capacity of transformed cancer cells that endogenously express the antigen; cells treated with specific 158P1D7 siRNAs show reduced survival as measured, e.g., by a metabolic readout of cell viability, correlating to the reduced proliferative capacity. Thus, 158P1D7 siRNA compositions comprise siRNA (double stranded RNA) that correspond to the nucleic acid ORF sequence of the 158P1D7 protein or subsequences thereof; these subsequences are generally 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30,31, 32, 33, 34, 35 or more than 35 contiguous RNA nucleotides in length and contain sequences that are complementary and non-complementary to at least a portion of the mRNA coding sequence In a preferred embodiment, the subsequences are 19-25 nucleotides in length, most preferably 21-23 nucleotides in length.

RNA interference is a novel approach to silencing genes *in vitro* and *in vivo,* thus small double stranded RNAs (siRNAs) are valuable therapeutic agents. The power of siRNAs to silence specific gene activities has now been brought to animal models of disease and is used in humans as well. For example, hydrodynamic infusion of a solution of siRNA into a mouse with a siRNA against a particular target has been proven to be therapeutically effective.

The pioneering work by Song *et al.* indicates that one type of entirely natural nucleic acid, small interfering RNAs (siRNAs), served as therapeutic agents even without further chemical modification (Song, E., et al. "RNA interference targeting Fas protects mice from fulminant hepatitis" Nat. Med. 9(3): 347-51(2003)). This work provided the first *in vivo* evidence that infusion of siRNAs into an animal could alleviate disease. In that case, the authors gave mice injections of siRNA designed to silence the FAS protein (a cell death receptor that when over-activated during inflammatory response induces hepatocytes and other cells to die). The next day, the animals were given an antibody specific to Fas. Control mice died of acute liver failure within a few days, while over 80% of the siRNA-treated mice remained free from serious disease and survived. About 80% to 90% of their liver cells incorporated the naked siRNA oligonucleotides. Furthermore, the RNA molecules functioned for 10 days before losing effect after 3 weeks.

For use in human therapy, siRNA is delivered by efficient systems that induce long-lasting RNAi activity. A major caveat for clinical use is delivering siRNAs to the appropriate cells. Hepatocytes seem to be particularly receptive to exogenous RNA. Today, targets located in the liver are attractive because liver is an organ that can be readily targeted by nucleic acid molecules and viral vectors. However, other tissue and organs targets are preferred as well.

Formulations of siRNAs with compounds that promote transit across cell membranes are used to improve administration of siRNAs in therapy. Chemically modified synthetic siRNA, that are resistant to nucleases and have serum stability have concomitant enhanced duration of RNAi effects, are an additional embodiment.

Thus, siRNA technology is a therapeutic for human malignancy by delivery of siRNA molecules directed to 158P1D7 to individuals with the cancers, such as those listed in Table 1. Such administration of siRNAs leads to reduced growth of cancer cells expressing 158P1D7, and provides an anti-tumor therapy, lessening the morbidity and/or mortality associated with malignancy.

The effectiveness of this modality of gene product knockdown is significant when measured *in vitro* or *in vivo.* Effectiveness *in vitro* is readily demonstrable through application of siRNAs to cells in culture (as described above) or to aliquots of cancer patient biopsies when *in vitro* methods are used to detect the reduced expression of 158P1D7 protein.

XV.) Kits/Articles of Manufacture

For use in the laboratory, prognostic, prophylactic, diagnostic and therapeutic applications described herein, kits are within the scope of the invention. Such kits can comprise a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements to be used in the method, along with a label or insert comprising instructions for use, such as a use described herein. For example, the container(s) can comprise a probe that is or can be detectably labeled. Such probe can be an antibody or polynucleotide specific for a protein or a gene or message of the invention, respectively. Where the method utilizes nucleic acid hybridization to detect the target nucleic acid, the kit can also have containers containing nucleotide(s) for amplification of the target nucleic acid sequence. Kits can comprise a container comprising a reporter, such as a biotin-binding protein, such as avidin or streptavidin, bound to a reporter molecule, such as an enzymatic, fluorescent, or radioisotope label; such a reporter can be used with, e.g., a nucleic acid or antibody. The kit can include all or part of the amino acid sequences in Figure 2 or Figure 3 or analogs thereof, or a nucleic acid molecule that encodes such amino acid sequences.

The kit of the invention will typically comprise the container described above and one or more other containers associated therewith that comprise materials desirable from a commercial and user standpoint, including buffers, diluents, filters, needles, syringes; carrier, package, container, vial and/or tube labels listing contents and/or instructions for use, and package inserts with instructions for use.

A label can be present on or with the container to indicate that the composition is used for a specific therapy or non-therapeutic application, such as a prognostic, prophylactic, diagnostic or laboratory application, and can also indicate directions for either *in vivo* or *in vitro* use, such as those described herein. Directions and or other information can also be included on an insert(s) or label(s) which is included with or on the kit. The label can be on or associated with the container. A label a can be on a container when letters, numbers or other characters forming the label are molded or etched into the container itself; a label can be associated with a container when it is present within a receptacle or carrier that also holds the container, e.g., as a package insert. The label can indicate that the composition is used for diagnosing, treating, prophylaxing or prognosing a condition, such as a neoplasia of a tissue set forth in Table I.

The terms "kit" and "article of manufacture" can be used as synonyms.

In another embodiment of the invention, an article(s) of manufacture containing compositions, such as amino acid sequence(s), small molecule(s), nucleic acid sequence(s), and/or antibody(s), e.g., materials useful for the diagnosis, prognosis, prophylaxis and/or treatment of neoplasias of tissues such as those set forth in Table I is provided. The article of manufacture typically comprises at least one container and at least one label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers can be formed from a variety of materials such as glass, metal or plastic. The container can hold amino acid sequence(s), small molecule(s), nucleic acid sequence(s), cell population(s) and/or antibody(s). In one embodiment, the container holds a polynucleotide for use in examining the mRNA expression profile of a cell, together with reagents used for this purpose. In another embodiment a container comprises an antibody, binding fragment thereof or specific binding protein for use in evaluating protein expression of 158P1D7 in cells and tissues, or for relevant laboratory, prognostic, diagnostic, prophylactic and therapeutic purposes; indications and/or directions for such uses can be included on or with such container, as can reagents and other compositions or tools used for these purposes. In another embodiment, a container comprises materials for eliciting a cellular or humoral immune response, together with associated indications and/or directions. In another embodiment, a container comprises materials for adoptive immunotherapy, such as cytotoxic T cells (CTL) or helper T cells (HTL), together with associated indications and/or directions; reagents and other compositions or tools used for such purpose can also be included.

The container can alternatively hold a composition that is effective for treating, diagnosis, prognosing or prophylaxing a condition and can have a sterile access port (for example the container can be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The active agents in the composition can be an antibody capable of specifically binding 158P1D7 and modulating the function of 158P1D7.

The article of manufacture can further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution and/or dextrose solution. It can further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, stirrers, needles, syringes, and/or package inserts with indications and/or instructions for use.

EXAMPLES

Various aspects of the invention are further described and illustrated by way of the several examples that follow, none of which are intended to limit the scope of the invention.

Example 1: SSH-Generated Isolation of a cDNA Fragment of the 158P1D7 Gene

To isolate genes that are over-expressed in bladder cancer we used the Suppression Subtractive Hybridization (SSH) procedure using cDNA derived from bladder cancer tissues, including invasive transitional cell carcinoma. The 158P1D7 SSH cDNA sequence was derived from a bladder cancer pool minus normal bladder cDNA subtraction. Included in the driver were also cDNAs derived from 9 other normal tissues. The 158P1D7 cDNA was identified as highly expressed in the bladder cancer tissue pool, with lower expression seen in a restricted set of normal tissues.

The SSH DNA sequence of 231 bp (Figure 1) has high homology (230/231 identity) to a hypothetical protein FLJ22774 (GenBank accession XM_033183) derived from a chromosome 13 genomic clone. A 158P1D7 cDNA clone (TurboScript3PX) of 2,555 bp was isolated from bladder cancer cDNA, revealing an ORF of 841 amino acids (Figure 2 and Figure 3).

The 158P1D7 protein has a signal sequence and a transmembrane domain and is predicted to be localized to the cell surface using the PSORT-I program (URL psort.nibb.ac.jp:8800/form.html). Amino acid sequence analysis of 158P1D7 reveals 100% identity over 798 amino acid region to a human hypothetical protein FLJ22774 (GenBank Accession XP_033182)(Figure 4).

Materials and Methods

Human Tissues:

The bladder cancer patient tissues were purchased from several sources such as from the NDRI (Philadelphia, PA). mRNA for some normal tissues were purchased from Clontech, Palo Alto, CA.

RNA Isolation:

Tissues were homogenized in Trizol reagent (Life Technologies, Gibco BRL) using 10 ml/ g tissue isolate total RNA. Poly A RNA was purified from total RNA using Qiagen's Oligotex mRNA Mini and Midi kits. Total and mRNA were quantified by spectrophotometric analysis (O.D. 260/280 nm) and analyzed by gel electrophoresis.

Oligonucleotides:

The following HPLC purified oligonucleotides were used.

DPNCDN (cDNA synthesis primer):

5'TTTT**GATC**AAGCTT$_{30}$3' (SEQ ID NO: 28)

Adaptor 1:

5'CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAG3' (SEQ ID NO: 29)

3'GGCCCGTC**CTAG**5' (SEQ ID NO: 30)

Adaptor 2:

5'GTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAG3' (SEQ ID NO: 31)

3'CGGCTC**CTAG**5' (SEQ ID NO: 32)

PCR primer 1:

5'CTAATACGACTCACTATAGGGC3' (SEQ ID NO: 33)

Nested primer (NP)1:

5'TCGAGCGGCCGCCCGGGCAG**GA**3' (SEQ ID NO: 34)

Nested primer (NP)2:

5'AGCGTGGTCGCGGCCGAG**GA**3' (SEQ ID NO: 35)

Suppression Subtractive Hybridization:

Suppression Subtractive Hybridization (SSH) was used to identify cDNAs corresponding to genes that may be differentially expressed in bladder cancer. The SSH reaction utilized cDNA from bladder cancer and normal tissues.

The gene 158P1D7 sequence was derived from a bladder cancer pool minus normal bladder cDNA subtraction. The SSH DNA sequence (Figure 1) was identified.

The cDNA derived from of pool of normal bladder tissues was used as the source of the "driver" cDNA, while the cDNA from a pool of bladder cancer tissues was used as the source of the "tester" cDNA. Double stranded cDNAs corresponding to tester and driver cDNAs were synthesized from 2 μg of poly(A)$^+$ RNA isolated from the relevant xenograft tissue, as described above, using CLONTECH's PCR-Select cDNA Subtraction Kit and 1 ng of oligonucleotide DPNCDN as primer. First- and second-strand synthesis were carried out as described in the Kit's user manual protocol (CLONTECH Protocol No. PT1117-1, Catalog No. K1804-1). The resulting cDNA was digested with Dpn II for 3 hrs at 37°C. Digested cDNA was extracted with phenol/chloroform (1:1) and ethanol precipitated.

Driver cDNA was generated by combining in a 1:1 ratio Dpn II digested cDNA from the relevant tissue source (see above) with a mix of digested cDNAs derived from the nine normal tissues: stomach, skeletal muscle, lung, brain, liver, kidney, pancreas, small intestine, and heart.

Tester cDNA was generated by diluting 1 μl of Dpn II digested cDNA from the relevant tissue source (see above) (400 ng) in 5 μl of water. The diluted cDNA (2 μl, 160 ng) was then ligated to 2 μl of Adaptor 1 and Adaptor 2 (10 μM), in separate ligation reactions, in a total volume of 10 μl at 16°C overnight, using 400 u of T4 DNA ligase (CLONTECH). Ligation was terminated with 1 μl of 0.2 M EDTA and heating at 72°C for 5 min.

The first hybridization was performed by adding 1.5 μl (600 ng) of driver cDNA to each of two tubes containing 1.5 μl (20 ng) Adaptor 1- and Adaptor 2- ligated tester cDNA. In a final volume of 4 μl, the samples were overlaid with mineral oil, denatured in an MJ Research thermal cycler at 98°C for 1.5 minutes, and then were allowed to hybridize for 8 hrs at

68°C. The two hybridizations were then mixed together with an additional 1 μl of fresh denatured driver cDNA and were allowed to hybridize overnight at 68°C. The second hybridization was then diluted in 200 μl of 20 mM Hepes, pH 8.3, 50 mM NaCl, 0.2 mM EDTA, heated at 70°C for 7 min. and stored at -20°C.

PCR Amplification, Cloning and Sequencing of Gene Fragments Generated from SSH:

To amplify gene fragments resulting from SSH reactions, two PCR amplifications were performed. In the primary PCR reaction 1 μl of the diluted final hybridization mix was added to 1 μl of PCR primer 1 (10 μM), 0.5 μl dNTP mix (10 μM), 2.5 μl 10 x reaction buffer (CLONTECH) and 0.5 μl 50 x Advantage cDNA polymerase Mix (CLONTECH) in a final volume of 25 μl. PCR 1 was conducted using the following conditions: 75°C for 5 min., 94°C for 25 sec., then 27 cycles of 94°C for 10 sec, 66°C for 30 sec, 72°C for 1.5 min. Five separate primary PCR reactions were performed for each experiment. The products were pooled and diluted 1:10 with water. For the secondary PCR reaction, 1 μl from the pooled and diluted primary PCR reaction was added to the same reaction mix as used for PCR 1, except that primers NP1 and NP2 (10 μM) were used instead of PCR primer 1. PCR 2 was performed using 10-12 cycles of 94°C for 10 sec, 68°C for 30 sec, and 72°C for 1.5 minutes. The PCR products were analyzed using 2% agarose gel electrophoresis.

The PCR products were inserted into pCR2.1 using the T/A vector cloning kit (Invitrogen). Transformed *E. coli* were subjected to blue/white and ampicillin selection. White colonies were picked and arrayed into 96 well plates and were grown in liquid culture overnight. To identify inserts, PCR amplification was performed on 1 ml of bacterial culture using the conditions of PCR1 and NP1 and NP2 as primers. PCR products were analyzed using 2% agarose gel electrophoresis.

Bacterial clones were stored in 20% glycerol in a 96 well format. Plasmid DNA was prepared, sequenced, and subjected to nucleic acid homology searches of the GenBank, dBest, and NCI-CGAP databases.

RT-PCR Expression Analysis:

First strand cDNAs can be generated from 1 μg of mRNA with oligo (dT)12-18 priming using the Gibco-BRL Superscript Preamplification system. The manufacturer's protocol was used which included an incubation for 50 min at 42°C with reverse transcriptase followed by RNAse H treatment at 37°C for 20 min. After completing the reaction, the volume can be increased to 200 μl with water prior to normalization. First strand cDNAs from 16 different normal human tissues can be obtained from Clontech.

Normalization of the first strand cDNAs from multiple tissues was performed by using the primers 5'atatcgccgcgctcgtcgtcgacaa3' (SEQ ID NO: 36) and 5'agccacacgcagctcattgtagaagg 3' (SEQ ID NO: 37) to amplify β-actin. First strand cDNA (5 μl) were amplified in a total volume of 50 μl containing 0.4 μM primers, 0.2 μM each dNTPs, 1XPCR buffer (Clontech, 10 mM Tris-HCL, 1.5 mM $MgCl_2$, 50 mM KCl, pH8.3) and 1X Klentaq DNA polymerase (Clontech). Five μl of the PCR reaction can be removed at 18, 20, and 22 cycles and used for agarose gel electrophoresis. PCR was performed using an MJ Research thermal cycler under the following conditions: Initial denaturation can be at 94°C for 15 sec, followed by a 18, 20, and 22 cycles of 94°C for 15, 65°C for 2 min, 72°C for 5 sec. A final extension at 72°C was carried out for 2 min. After agarose gel electrophoresis, the band intensities of the 283 b.p. β-actin bands from multiple tissues were compared by visual inspection. Dilution factors for the first strand cDNAs were calculated to result in equal β-actin band intensities in all tissues after 22 cycles of PCR. Three rounds of normalization can be required to achieve equal band intensities in all tissues after 22 cycles of PCR.

To determine expression levels of the 158P1D7 gene, 5 μl of normalized first strand cDNA were analyzed by PCR using 26, and 30 cycles of amplification. Semi-quantitative expression analysis can be achieved by comparing the PCR products at cycle numbers that give light band intensities. The primers used for RT-PCR were designed using the 158P1D7 SSH sequence and are listed below:

158P1D7.1

5' ATAAGCTTTCAATGTTGCGCTCCT 3' (SEQ ID NO: 38)

158P1D7.2

5' TGTCAACTAAGACCACGTCCATTC3' (SEQ ID NO: 39)

A typical RT-PCR expression analysis is shown in Figure 6. RT-PCR expression analysis was performed on first strand cDNAs generated using pools of tissues from multiple samples. The cDNAs were shown to be normalized using beta-actin PCR. Expression of 158P1D7 was observed in bladder cancer pool.

Example 2: Full Length Cloning of 158P1D7

The 158P1D7 SSH cDNA sequence was derived from a bladder cancer pool minus normal bladder cDNA subtraction. The SSH cDNA sequence (Figure 1) was designated 158P1D7. The full-length cDNA clone 158P1D7-clone TurboScript3PX (Figure 2) was cloned from bladder cancer pool cDNA.

158P1D7 clone cDNA was deposited under the terms of the Budapest Treaty on 22 August 2001, with the American Type Culture Collection (ATCC; 10801 University Blvd., Manassas, VA 20110-2209 USA) as plasmid p158P1D7-Turbo/3PX, and has been assigned Accession No. PTA-3662.

Example 3: Chromosomal Mapping of 158P1D7

Chromosomal localization can implicate genes in disease pathogenesis. Several chromosome mapping approaches are available including fluorescent *in situ* hybridization (FISH), human/hamster radiation hybrid (RH) panels (Walter et al., 1994; Nature Genetics 7:22; Research Genetics, Huntsville Al), human-rodent somatic cell hybrid panels such as is available from the Coriell Institute (Camden, New Jersey), and genomic viewers utilizing BLAST homologies to sequenced and mapped genomic clones (NCBI, Bethesda, Maryland).

158P1D7 maps to chromosme 13, using 158P1D7 sequence and the NCBI BLAST tool: (world wide web URL ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs). This is a region of frequent amplification in bladder cancer (Prat et al., Urology 2001 May;57(5):986-92; Muscheck et al., Carcinogenesis 2000 Sep;21(9):1721-26) and is associated with rapid tumor cell proliferation in advanced bladder cancer (Tomovska et al., Int J Oncol 2001 Jun;18(6):1239-44).

Example 4: Expression analysis of 158P1D7 in normal tissues and patient specimens

Analysis of 158P1D7 by RT-PCR is shown in Figure 6. Strong expression of 158P1D7 is observed in bladder cancer pool and breast cancer pool. Lower levels of expression are observed in VP1, VP2, xenograft pool, prostate cancer pool, colon cancer pool , lung cancer pool, ovary cancer pool, and metastasis pool.

Extensive northern blot analysis of 158P1D7 in 16 human normal tissues confirms the expression observed by RT-PCR (Figure 7). Two transcripts of approximately 4.6 and 4.2 kb are detected in prostate and, to lower levels, in heart, placenta, liver, small intestine and colon.

Northern blot analysis on patient tumor specimens shows expression of 158P1D7 in most bladder tumor tissues tested and in the bladder cancer cell line SCaBER (Figure 8A and 8B). The expression detected in normal adjacent tissues (isolated from patients) but not in normal tissues (isolated from a healthy donor) may indicate that these tissues are not fully normal and that 158P1D7 may be expressed in early stage tumors. Expression of 158P1D7 is also detected in 2 of 4 lung cancer cell lines, and in all 3 lung cancer tissues tested (Figure 9). In breast cancer samples, 158P1D7 expression is

observed in the MCF7 and CAMA-1 breast cancer cell lines, in breast tumor tissues isolated from breast cancer patients, but not in normal breast tissues (Figure 10). 158P1D7 shows expression in melanoma cancer. RNA was extracted from normal skin cell line Detroit-551, and from the melanoma cancer cell line A375. Northern blots with 10ug of total RNA were probed with the 158P1D7 DNA probe. Results show expression of 158P1D7 in the melanoma cancer cell line but not in the normal cell line (Figure 20). 158P1D7 shows expression in cervical cancer patient specimens. First strand cDNA was prepared from normal cervix, cervical cancer cell line Hela, and a panel of cervical cancer patient specimens. Normalization was performed by PCR using primers to actin and GAPDH. Semi-quantitative PCR, using primers to 158P1D7, was performed at 26 and 30 cycles of amplification. Results show expression of 158P1D7 in 5 out of 14 tumor specimens tested but not in normal cervix nor in the cell line (Figure 21).

The restricted expression of 158P1D7 in normal tissues and the expression detected in prostate cancer, bladder cancer, colon cancer, lung cancer, ovarian cancer, breast cancer, melanoma cancer, and cervical cancer suggest that 158P1D7 is a potential therapeutic target and a diagnostic marker for human cancers.

Example 5: Production of Recombinant 158P1D7 in Prokaryotic Systems

To express recombinant 158P1D7 and 158P1D7 variants in prokaryotic cells, the full or partial length 158P1D7 and 158P1D7 variant cDNA sequences are cloned into any one of a variety of expression vectors known in the art. One or more of the following regions of 158P1D7 variants are expressed: the full length sequence presented in Figures 2 and 3, or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 158P1D7, variants, or analogs thereof.

A. *In vitro* transcription and translation constructs:

pCRII: To generate 158P1D7 sense and anti-sense RNA probes for RNA *in situ* investigations, pCRII constructs (Invitrogen, Carlsbad CA) are generated encoding either all or fragments of the 158P1D7 cDNA. The pCRII vector has Sp6 and T7 promoters flanking the insert to drive the transcription of 158P1D7 RNA for use as probes in RNA *in situ* hybridization experiments. These probes are used to analyze the cell and tissue expression of 158P1D7 at the RNA level. Transcribed 158P1D7 RNA representing the cDNA amino acid coding region of the 158P1D7 gene is used in *in vitro* translation systems such as the TnT™ Coupled Reticulolysate System (Promega, Corp., Madison, WI) to synthesize 158P1D7 protein.

B. Bacterial Constructs:

pGEX Constructs: To generate recombinant 158P1D7 proteins in bacteria that are fused to the Glutathione S-transferase (GST) protein, all or parts of the 158P1D7 cDNA protein coding sequence are cloned into the pGEX family of GST-fusion vectors (Amersham Pharmacia Biotech, Piscataway, NJ). These constructs allow controlled expression of recombinant 158P1D7 protein sequences with GST fused at the amino-terminus and a six histidine epitope (6X His) at the carboxyl-terminus. The GST and 6X His tags permit purification of the recombinant fusion protein from induced bacteria with the appropriate affinity matrix and allow recognition of the fusion protein with anti-GST and anti-His antibodies. The 6X His tag is generated by adding 6 histidine codons to the cloning primer at the 3' end, e.g., of the open reading frame (ORF). A proteolytic cleavage site, such as the PreScission™ recognition site in pGEX-6P-1, may be employed such that it permits cleavage of the GST tag from 158P1D7-related protein. The ampicillin resistance gene and pBR322 origin permits selection and maintenance of the pGEX plasmids in *E. coli*.

pMAL Constructs: To generate, in bacteria, recombinant 158P1D7 proteins that are fused to maltose-binding protein (MBP), all or parts of the 158P1D7 cDNA protein coding sequence are fused to the MBP gene by cloning into the pMAL-c2X and pMAL-p2X vectors (New England Biolabs, Beverly, MA). These constructs allow controlled expression of recombinant 158P1D7 protein sequences with MBP fused at the amino-terminus and a 6X His epitope tag at the carboxyl-terminus. The MBP and 6X His tags permit purification of the recombinant protein from induced bacteria with the appropriate

affinity matrix and allow recognition of the fusion protein with anti-MBP and anti-His antibodies. The 6X His epitope tag is generated by adding 6 histidine codons to the 3' cloning primer. A Factor Xa recognition site permits cleavage of the pMAL tag from 158P1D7. The pMAL-c2X and pMAL-p2X vectors are optimized to express the recombinant protein in the cytoplasm or periplasm respectively. Periplasm expression enhances folding of proteins with disulfide bonds. Amino acids 356-608 of 158P1D7 variant 1 have been cloned into the pMALc2X vector.

pET Constructs: To express 158P1D7 in bacterial cells, all or parts of the 158P1D7 cDNA protein coding sequence are cloned into the pET family of vectors (Novagen, Madison, WI). These vectors allow tightly controlled expression of recombinant 158P1D7 protein in bacteria with and without fusion to proteins that enhance solubility, such as NusA and thioredoxin (Trx), and epitope tags, such as 6X His and S-Tag ™ that aid purification and detection of the recombinant protein. For example, constructs are made utilizing pET NusA fusion system 43.1 such that regions of the 158P1D7 protein are expressed as amino-terminal fusions to NusA.

C. Yeast Constructs:

pESC Constructs: To express 158P1D7 in the yeast species *Saccharomyces cerevisiae* for generation of recombinant protein and functional studies, all or parts of the 158P1D7 cDNA protein coding sequence are cloned into the pESC family of vectors each of which contain 1 of 4 selectable markers, HIS3, TRP1, LEU2, and URA3 (Stratagene, La Jolla, CA). These vectors allow controlled expression from the same plasmid of up to 2 different genes or cloned sequences containing either Flag™ or Myc epitope tags in the same yeast cell. This system is useful to confirm protein-protein interactions of 158P1D7. In addition, expression in yeast yields similar post-translational modifications, such as glycosylations and phosphorylations, that are found when expressed in eukaryotic cells.

pESP Constructs: To express 158P1D7 in the yeast species *Saccharomyces pombe*, all or parts of the 158P1D7 cDNA protein coding sequence are cloned into the pESP family of vectors. These vectors allow controlled high level of expression of a 158P1D7 protein sequence that is fused at either the amino terminus or at the carboxyl terminus to GST which aids purification of the recombinant protein. A Flag™ epitope tag allows detection of the recombinant protein with anti-Flag™ antibody.

Example 6: Production of Recombinant 158P1D7 in Eukaryotic Systems

A. Mammalian Constructs:

To express recombinant 158P1D7 in eukaryotic cells, the full or partial length 158P1D7 cDNA sequences were cloned into any one of a variety of expression vectors known in the art. One or more of the following regions of 158P1D7 were expressed in these constructs, amino acids 1 to 841, or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 158P1D7 v.1; amino acids 1 to 732 of v.3; amino acids 1 to 395 of v.4; amino acids 1 to 529 of v.6; or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 158P1D7 variants, or analogs thereof.

The constructs can be transfected into any one of a wide variety of mammalian cells such as 293T cells. Transfected 293T cell lysates can be probed with the anti-158P1D7 polyclonal serum, described herein.

pcDNA4/HisMax Constructs: To express 158P1D7 in mammalian cells, a 158P1D7 ORF, or portions thereof, of 158P1D7 are cloned into pcDNA4/HisMax Version A (Invitrogen, Carlsbad, CA). Protein expression is driven from the cytomegalovirus (CMV) promoter and the SP16 translational enhancer. The recombinant protein has Xpress™ and six histidine (6X His) epitopes fused to the amino-terminus. The pcDNA4/HisMax vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Zeocin

resistance gene allows for selection of mammalian cells expressing the protein and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli.*

pcDNA3.1/MycHis Constructs: To express 158P1D7 in mammalian cells, a 158P1D7 ORF, or portions thereof, of 158P1D7 with a consensus Kozak translation initiation site was cloned into pcDNA3.1/MycHis Version A (Invitrogen, Carlsbad, CA). Protein expression was driven from the cytomegalovirus (CMV) promoter. The recombinant proteins have the myc epitope and 6X His epitope fused to the carboxyl-terminus. The pcDNA3.1/MycHis vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability, along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Neomycin resistance gene can be used, as it allows for selection of mammalian cells expressing the protein and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli.*

The complete ORF of 158P1D7 v.1 was cloned into the pcDNA3.1/MycHis construct to generate 158P1D7.pcDNA3.1/MycHis. Figure 23 shows expression of 158P1D7.pcDNA3.1/MycHis following transfection into 293T cells. 293T cells were transfected with either 158P1D7.pcDNA3.1/MycHis or pcDNA3.1/MycHis vector control. Forty hours later, cells were collected and analyzed by flow cytometry using anti-158P1D7 monoclonal antibodies. Results show expression of 158P1D7 from the 158P1D7.pcDNA3.1/MycHis construct on the surface of transfected cells.

pcDNA3.1/CT-GFP-TOPO Construct: To express 158P1D7 in mammalian cells and to allow detection of the recombinant proteins using fluorescence, a 158P1D7 ORF, or portions thereof, with a consensus Kozak translation initiation site are cloned into pcDNA3.1/CT-GFP-TOPO (Invitrogen, CA). Protein expression is driven from the cytomegalovirus (CMV) promoter. The recombinant proteins have the Green Fluorescent Protein (GFP) fused to the carboxyl-terminus facilitating non-invasive, *in vivo* detection and cell biology studies. The pcDNA3.1CT-GFP-TOPO vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Neomycin resistance gene allows for selection of mammalian cells that express the protein, and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli.* Additional constructs with an amino-terminal GFP fusion are made in pcDNA3.1/NT-GFP-TOPO spanning the entire length of a 158P1D7 protein.

PAPtag: A 158P1D7 ORF, or portions thereof, is cloned into pAPtag-5 (GenHunter Corp. Nashville, TN). This construct generates an alkaline phosphatase fusion at the carboxyl-terminus of a 158P1D7 protein while fusing the IgGκ signal sequence to the amino-terminus. Constructs are also generated in which alkaline phosphatase with an amino-terminal IgGκ signal sequence is fused to the amino-terminus of a 158P1D7 protein. The resulting recombinant 158P1D7 proteins are optimized for secretion into the media of transfected mammalian cells and can be used to identify proteins such as ligands or receptors that interact with 158P1D7 proteins. Protein expression is driven from the CMV promoter and the recombinant proteins also contain myc and 6X His epitopes fused at the carboxyl-terminus that facilitates detection and purification. The Zeocin resistance gene present in the vector allows for selection of mammalian cells expressing the recombinant protein and the ampicillin resistance gene permits selection of the plasmid in *E. coli.*

pTag5: A 158P1D7 ORF, or portions thereof, were cloned into pTag-5. This vector is similar to pAPtag but without the alkaline phosphatase fusion. This construct generated a 158P1D7 protein with an amino-terminal IgGκ signal sequence and myc and 6X His epitope tags at the carboxyl-terminus that facilitate detection and affinity purification. The resulting recombinant 158P1D7 protein was optimized for secretion into the media of transfected mammalian cells, and was used as immunogen or ligand to identify proteins such as ligands or receptors that interact with the 158P1D7 proteins. Protein expression is driven from the CMV promoter. The Zeocin resistance gene present in the vector allows for selection of mammalian cells expressing the protein, and the ampicillin resistance gene permits selection of the plasmid in *E. coli.*

The extracellular domain, amino acids 16-608, 27-300, and 301-608, of 158P1D7 v.1 were cloned into the pTag5 construct to generate 158P1D7(16-608).pTag5, 158P1D7(27-300).pTag5, and 158P1D7(301-608).pTag5 respectively. Expression and secretion of the various segments of the extracellular domain of 158P1D7 following vector transfection into 293T cells was confirmed.

PsecFc: A 158P1D7 ORF, or portions thereof, was also cloned into psecFc. The psecFc vector was assembled by cloning the human immunoglobulin G1 (IgG) Fc (hinge, CH2, CH3 regions) into pSecTag2 (Invitrogen, California). This construct generates an IgG1 Fc fusion at the carboxyl-terminus of the 158P1D7 proteins, while fusing the IgGK signal sequence to N-terminus. 158P1D7 fusions utilizing the murine IgG1 Fc region are also used. The resulting recombinant 158P1D7 proteins are optimized for secretion into the media of transfected mammalian cells, and can be used as immunogens or to identify proteins such as ligands or receptors that interact with 158P1D7 protein. Protein expression is driven from the CMV promoter. The hygromycin resistance gene present in the vector allows for selection of mammalian cells that express the recombinant protein, and the ampicillin resistance gene permits selection of the plasmid in *E. coli*.

The extracellular domain amino acids 16-608 of 158P1D7 v.1 was cloned into the psecFc construct to generate 158P1D7(16-608).psecFc.

pSRα Constructs: To generate mammalian cell lines that express 158P1D7 constitutively, 158P1D7 ORF, or portions thereof, of 158P1D7 were cloned into pSRα constructs. Amphotropic and ecotropic retroviruses were generated by transfection of pSRα constructs into the 293T-10A1 packaging line or co-transfection of pSRα and a helper plasmid (containing deleted packaging sequences) into the 293 cells, respectively. The retrovirus is used to infect a variety of mammalian cell lines, resulting in the integration of the cloned gene, 158P1D7, into the host cell-lines. Protein expression is driven from a long terminal repeat (LTR). The Neomycin resistance gene present in the vector allows for selection of mammalian cells that express the protein, and the ampicillin resistance gene and ColE1 origin permit selection and maintenance of the plasmid in *E. coli*. The retroviral vectors can thereafter be used for infection and generation of various cell lines using, for example, PC3, NIH 3T3, TsuPr1, 293 or rat-1 cells.

The complete ORF of 158P1D7 v.1 was cloned into the pSRα construct to generate 158P1D7.pSRα. Figure 23 shows expression of 158P1D7.pSRα following trasnduction into UMUC3 cells. UMUC-3 cells were transduced with either 158P1D7.pSRα or vector control. Forty hours later, cells were collected and analyzed by flow cytometry using anti-158P1D7 monoclonal antibodies. Results show expression of 158P1D7 from the 158P1D7.pSRα construct on the surface of the cells.

Additional pSRα constructs are made that fuse an epitope tag such as the FLAG™ tag to the carboxyl-terminus of 158P1D7 sequences to allow detection using anti-Flag antibodies. For example, the FLAG™ sequence 5' gat tac aag gat gac gac gat aag 3' (SEQ ID NO: 40) is added to cloning primer at the 3' end of the ORF. Additional pSRα constructs are made to produce both amino-terminal and carboxyl-terminal GFP and myc/6X His fusion proteins of the full-length 158P1D7 proteins.

Additional Viral Vectors: Additional constructs are made for viral-mediated delivery and expression of 158P1D7. High virus titer leading to high level expression of 158P1D7 is achieved in viral delivery systems such as adenoviral vectors and herpes amplicon vectors. A 158P1D7 coding sequences or fragments thereof are amplified by PCR and subcloned into the AdEasy shuttle vector (Stratagene). Recombination and virus packaging are performed according to the manufacturer's instructions to generate adenoviral vectors. Alternatively, 158P1D7 coding sequences or fragments thereof are cloned into the HSV-1 vector (Imgenex) to generate herpes viral vectors. The viral vectors are thereafter used for infection of various cell lines such as PC3, NIH 3T3, 293 or rat-1 cells.

Regulated Expression Systems: To control expression of 158P1D7 in mammalian cells, coding sequences of 158P1D7, or portions thereof, are cloned into regulated mammalian expression systems such as the T-Rex System

(Invitrogen), the GeneSwitch System (Invitrogen) and the tightly-regulated Ecdysone System (Sratagene). These systems allow the study of the temporal and concentration dependent effects of recombinant 158P1D7. These vectors are thereafter used to control expression of 158P1D7 in various cell lines such as PC3, NIH 3T3, 293 or rat-1 cells.

B. Baculovirus Expression Systems

To generate recombinant 158P1D7 proteins in a baculovirus expression system, 158P1D7 ORF, or portions thereof, are cloned into the baculovirus transfer vector pBlueBac 4.5 (Invitrogen), which provides a His-tag at the N-terminus. Specifically, pBlueBac-158P1D7 is co-transfected with helper plasmid pBac-N-Blue (Invitrogen) into SF9 (*Spodoptera frugiperda*) insect cells to generate recombinant baculovirus (see Invitrogen instruction manual for details). Baculovirus is then collected from cell supernatant and purified by plaque assay.

Recombinant 158P1D7 protein is then generated by infection of HighFive insect cells (Invitrogen) with purified baculovirus. Recombinant 158P1D7 protein can be detected using anti-158P1D7 or anti-His-tag antibody. 158P1D7 protein can be purified and used in various cell-based assays or as immunogen to generate polyclonal and monoclonal antibodies specific for 158P1D7.

Example 7 Antigenicity Profiles and Secondary Structure

Figure 11(a)-(d), Figure 12(a)-(d), Figure 13(a)-(d), Figure 14(a)-(d), and Figure 15(a)-(d) depict graphically five amino acid profiles each of 158P1D7 protein variants 1, 3, 4, and 6, each assessment available by accessing the ProtScale website located on the World Wide Web at (.expasy.ch/cgi-bin/protscale.pl) on the ExPasy molecular biology server.

These profiles: Figure 11, Hydrophilicity, (Hopp T.P., Woods K.R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828); Figure 12, Hydropathicity, (Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132); Figure 13, Percentage Accessible Residues (Janin J., 1979 Nature 277:491-492); Figure 14, Average Flexibility, (Bhaskaran R., and Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255); Figure 15, Beta-turn (Deleage, G., Roux B. 1987 Protein Engineering 1:289-294); and optionally others available in the art, such as on the ProtScale website, were used to identify antigenic regions of each of the 158P1D7 variant proteins. Each of the above amino acid profiles of 158P1D7 variants were generated using the following ProtScale parameters for analysis: 1) A window size of 9; 2) 100% weight of the window edges compared to the window center; and, 3) amino acid profile values normalized to lie between 0 and 1.

Hydrophilicity (Figure 11), Hydropathicity (Figure 12) and Percentage Accessible Residues (Figure 13) profiles were used to determine stretches of hydrophilic amino acids (i.e., values greater than 0.5 on the Hydrophilicity and Percentage Accessible Residues profile, and values less than 0.5 on the Hydropathicity profile). Such regions are likely to be exposed to the aqueous environment, be present on the surface of the protein, and thus available for immune recognition, such as by antibodies.

Average Flexibility (Figure 14) and Beta-turn (Figure 15) profiles determine stretches of amino acids (i.e., values greater than 0.5 on the Beta-turn profile and the Average Flexibility profile) that are not constrained in secondary structures such as beta sheets and alpha helices. Such regions are also more likely to be exposed on the protein and thus accessible to immune recognition, such as by antibodies.

Antigenic sequences of the 158P1D7 variant proteins indicated, e.g., by the profiles set forth in Figures 11(a)-(d), Figure 12(a)-(d), Figure 13(a)-(d), Figure 14(a)-(d), and Figure 15(a)-(d) are used to prepare immunogens, either peptides or nucleic acids that encode them, to generate therapeutic and diagnostic anti-158P1D7 antibodies. The immunogen can be any 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 or more than 50 contiguous amino acids, or the corresponding nucleic acids that encode them, from the 158P1D7 protein variants listed in Figures 2 and 3. In particular, peptide immunogens of the invention can comprise, a peptide region of at least 5 amino acids of Figures 2 and 3 in any whole number increment that includes an amino acid position having a value greater than 0.5 in the

Hydrophilicity profiles of Figure 11; a peptide region of at least 5 amino acids of Figures 2 and 3 in any whole number increment that includes an amino acid position having a value less than 0.5 in the Hydropathicity profile of Figures 12; a peptide region of at least 5 amino acids of Figures 2 and 3 in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Percent Accessible Residues profiles of Figure 13; a peptide region of at least 5 amino acids of Figures 2 and 3 in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Average Flexibility profiles on Figure 14; and, a peptide region of at least 5 amino acids of Figures 2 and 3 in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Beta-turn profile of Figures 15. Peptide immunogens of the invention can also comprise nucleic acids that encode any of the forgoing.

All immunogens of the invention, peptide or nucleic acid, can be embodied in human unit dose form, or comprised by a composition that includes a pharmaceutical excipient compatible with human physiology.

The secondary structure of 158P1D7 protein variants 1, 3, 4, and 6, namely the predicted presence and location of alpha helices, extended strands, and random coils, are predicted from the primary amino acid sequence using the HNN - Hierarchical Neural Network method (NPS@: Network Protein Sequence Analysis TIBS 2000 March Vol. 25, No 3 [291]:147-150 Combet C., Blanchet C., Geourjon C. and Deléage G., http://pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html), accessed from the ExPasy molecular biology server (http://www.expasy.ch/tools/). The analysis indicates that 158P1D7 variant 1 is composed of 35.32% alpha helix, 15.93% extended strand, and 48.75% random coil (Figure 16A). Variant 3 is composed of 34.97% alpha helix, 16.94% extended strand, and 48.09% random coil (Figure 16B). Variant 4 is composed of 24.56% alpha helix, 20.76% extended strand, and 54.68 % random coil (Figure 16C). Variant 6 is composed of 28.92% alpha helix, 17.96% extended strand, and 53.12% random coil (Figure 16D).

Analysis for the potential presence of transmembrane domains in the 158P1D7 variant proteins was carried out using a variety of transmembrane prediction algorithms accessed from the ExPasy molecular biology server (http://www.expasy.ch/tools/). Shown graphically in figure 16E, 16G, 16I, 16K, are the results of analysis of variants 1, 3, 4, and 6, respectively, using the TMpred program. In figure 16F, 16H, 16I, 16L are the results of variants 1, 3, 4, and 6, respectively, using the TMHMM program. Both the TMpred program and the TMHMM program predict the presence of 1 transmembrane domain in variant 1 and 3. Variants 4 and 6 are not predicted to contain transmembrane domains. All variants contain a stretch of hydrophobic amino acid sequence at their amino terminus that may encode a signal peptide. Analyses of 158P1D7 and 158P1D7 variants using other structural prediction programs are summarized in Table LVI.

Example 8: Generation of 158P1D7 Polyclonal Antibodies

Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. In addition to immunizing with a full length 158P1D7 protein variant, computer algorithms are employed in design of immunogens that, based on amino acid sequence analysis contain characteristics of being antigenic and available for recognition by the immune system of the immunized host (see the Example entitled "Antigenicity Profiles and Secondary Structure"). Such regions would be predicted to be hydrophilic, flexible, in beta-turn conformations, and be exposed on the surface of the protein (see, e.g., Figure 11, Figure 12, Figure 13, Figure 14, or Figure 15 for amino acid profiles that indicate such regions of 158P1D7 protein variants 1, 3, 4, and 6).

For example, recombinant bacterial fusion proteins or peptides containing hydrophilic, flexible, beta-turn regions of 158P1D7 protein variants are used as antigens to generate polyclonal antibodies in New Zealand White rabbits or monoclonal antibodies as described in Example 9. For example, in 158P1D7 variant 1, such regions include, but are not limited to, amino acids 25-45, amino acids 250-385, and amino acids 694-730. It is useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include,

but are not limited to, keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. In one embodiment, a peptide encoding amino acids 274-285 of 158P1D7 variant 1 was synthesized and conjugated to KLH. This peptide is then used as immunogen. Alternatively the immunizing agent may include all or portions of the 158P1D7 variant proteins, analogs or fusion proteins thereof. For example, the 158P1D7 variant 1 amino acid sequence can be fused using recombinant DNA techniques to any one of a variety of fusion protein partners that are well known in the art, such as glutathione-S-transferase (GST) and HIS tagged fusion proteins. In another embodiment, amino acids 27-300 of 158P1D7 variant 1 is fused to GST using recombinant techniques and the pGEX expression vector, expressed, purified and used to immunize a rabbit. Such fusion proteins are purified from induced bacteria using the appropriate affinity matrix.

Other recombinant bacterial fusion proteins that may be employed include maltose binding protein, LacZ, thioredoxin, NusA, or an immunoglobulin constant region (see the section entitled "Production of 158P1D7 in Prokaryotic Systems" and Current Protocols In Molecular Biology, Volume 2, Unit 16, Frederick M. Ausubul et al. eds., 1995; Linsley, P.S., Brady, W., Urnes, M., Grosmaire, L., Damle, N., and Ledbetter, L.(1991) J.Exp. Med. 174, 561-566).

In addition to bacterial derived fusion proteins, mammalian expressed protein antigens are also used. These antigens are expressed from mammalian expression vectors such as the Tag5 and Fc-fusion vectors (see the section entitled "Production of Recombinant 158P1D7 in Eukaryotic Systems"), and retain post-translational modifications such as glycosylations found in native protein. In one embodiment, amino acids 16-608 of 158P1D7 variant 1 was cloned into the Tag5 mammalian secretion vector, and expressed in 293T cells. The recombinant protein was purified by metal chelate chromatography from tissue culture supernatants of 293T cells stably expressing the recombinant vector. The purified Tag5 158P1D7 variant 1 protein is then used as immunogen.

During the immunization protocol, it is useful to mix or emulsify the antigen in adjuvants that enhance the immune response of the host animal. Examples of adjuvants include, but are not limited to, complete Freund's adjuvant (CFA) and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

In a typical protocol, rabbits are initially immunized subcutaneously with up to 200 μg, typically 100-200 μg, of fusion protein or peptide conjugated to KLH mixed in complete Freund's adjuvant (CFA). Rabbits are then injected subcutaneously every two weeks with up to 200 μg, typically 100-200 μg, of the immunogen in incomplete Freund's adjuvant (IFA). Test bleeds are taken approximately 7-10 days following each immunization and used to monitor the titer of the antiserum by ELISA.

To test reactivity and specificity of immune serum, such as the rabbit serum derived from immunization with the GST-fusion of 158P1D7 variant 1 protein, the full-length 158P1D7 variant 1 cDNA is cloned into pCDNA 3.1 myc-his expression vector (Invitrogen, see the Example entitled "Production of Recombinant 158P1D7 in Eukaryotic Systems"). After transfection of the constructs into 293T cells, cell lysates are probed with the anti-158P1D7 serum and with anti-His antibody (Santa Cruz Biotechnologies, Santa Cruz, CA) to determine specific reactivity to denatured 158P1D7 protein using the Western blot technique. In addition, the immune serum is tested by fluorescence microscopy, flow cytometry and immunoprecipitation against 293T and other recombinant 158P1D7-expressing cells to determine specific recognition of native protein. Western blot, immunoprecipitation, fluorescent microscopy, and flow cytometric techniques using cells that endogenously express 158P1D7 are also carried out to test reactivity and specificity.

Anti-serum from rabbits immunized with 158P1D7 variant fusion proteins, such as GST and MBP fusion proteins, are purified by depletion of antibodies reactive to the fusion partner sequence by passage over an affinity column containing the fusion partner either alone or in the context of an irrelevant fusion protein. For example, antiserum derived from a GST-158P1D7 variant 1 fusion protein is first purified by passage over a column of GST protein covalently coupled to AffiGel matrix (BioRad, Hercules, Calif.). The antiserum is then affinity purified by passage over a column composed of a MBP-

158P1D7 fusion protein covalently coupled to Affigel matrix. The serum is then further purified by protein G affinity chromatography to isolate the IgG fraction. Sera from other His-tagged antigens and peptide immunized rabbits as well as fusion partner depleted sera are affinity purified by passage over a column matrix composed of the original protein immunogen or free peptide.

Example 9: Generation of 158P1D7 Monoclonal Antibodies (mAbs)

In one embodiment, therapeutic mAbs to 158P1D7 variants comprise those that react with epitopes specific for each variant protein or specific to sequences in common between the variants that would bind, internalize, disrupt or modulate the biological function of the 158P1D7 variants, for example those that would disrupt the interaction with ligands and binding partners. Immunogens for generation of such mAbs include those designed to encode or contain the extracellular domain or the entire 158P1D7 protein variant sequence, regions predicted to contain functional motifs, and regions of the 158P1D7 protein variants predicted to be antigenic from computer analysis of the amino acid sequence (see, e.g., Figure 11, Figure 12, Figure 13, Figure 14, or Figure 15, and the Example entitled "Antigenicity Profiles and Secondary Structure"). Immunogens include peptides, recombinant bacterial proteins, and mammalian expressed Tag 5 proteins and human and murine IgG FC fusion proteins. In addition, pTAG5 protein, DNA vectors encoding the pTAG5 cells engineered to express high levels of a respective 158P1D7 variant, such as 293T-158P1D7 variant 1 or 3T3, RAT, or 300.19-158P1D7 variant 1murine Pre-B cells, are used to immunize mice.

To generate mAbs to a 158P1D7 variant, mice are first immunized intraperitoneally (IP) with, typically, 10-50 μg of protein immunogen or 10^7 158P1D7-expressing cells mixed in complete Freund's adjuvant. Mice are then subsequently immunized IP every 2-4 weeks with, typically, 10-50 μg of protein immunogen or 10^7 cells mixed in incomplete Freund's adjuvant. Alternatively, MPL-TDM adjuvant is used in immunizations. In addition to the above protein and cell-based immunization strategies, a DNA-based immunization protocol is employed in which a mammalian expression vector encoding a 158P1D7 variant sequence is used to immunize mice by direct injection of the plasmid DNA. For example, amino acids 16-608 of 158P1D7 of variant 1 was cloned into the Tag5 mammalian secretion vector and the recombinant vector was used as immunogen. In another example, the same amino acids were cloned into an Fc-fusion secretion vector in which the 158P1D7 variant 1 sequence is fused at the amino-terminus to an IgK leader sequence and at the carboxyl-terminus to the coding sequence of the human or murine IgG Fc region. This recombinant vector was then used as immunogen. The plasmid immunization protocols were used in combination with purified proteins expressed from the same vector and with cells expressing the respective 158P1D7 variant.

During the immunization protocol, test bleeds are taken 7-10 days following an injection to monitor titer and specificity of the immune response. Once appropriate reactivity and specificity is obtained as determined by ELISA, Western blotting, immunoprecipitation, fluorescence microscopy, and flow cytometric analyses, fusion and hybridoma generation is then carried out with established procedures well known in the art (see, e.g., Harlow and Lane, 1988).

In one embodiment for generating 158P1D7 variant 1 monoclonal antibodies, a peptide encoding amino acids 274-285 was synthesized, conjugated to KLH and used as immunogen. ELISA on free peptide was used to identify immunoreactive clones. Reactivity and specificity of the monoclonal antibodies to full length 158P1D7 variant 1 protein was monitored by Western blotting, immunoprecipitation, and flow cytometry using both recombinant and endogenous-expressing 158P1D7 variant 1 cells (See Figures 22, 23, 24, 25, and 28).

The binding affinity of 158P1D7 variant 1 specific monoclonal antibodies was determined using standard technologies. Affinity measurements quantify the strength of antibody to epitope binding and are used to help define which 158P1D7 variant monoclonal antibodies preferred for diagnostic or therapeutic use, as appreciated by one of skill in the art.

The BIAcore system (Uppsala, Sweden) is a preferred method for determining binding affinity. The BIAcore system uses surface plasmon resonance (SPR, Welford K. 1991, Opt. Quant. Elect. 23:1; Morton and Myszka, 1998, Methods in Enzymology 295: 268) to monitor biomolecular interactions in real time. BIAcore analysis conveniently generates association rate constants, dissociation rate constants, equilibrium dissociation constants, and affinity constants. Results of BIAcore analysis of 158P1D7 variant 1 monoclonal antibodies is shown in Table LVII.

To generate monoclonal antibodies specific for other 158P1D7 variants, immunogens are designed to encode amino acid sequences unique to the variants. In one embodiment, a peptide encoding amino acids 382-395 unique to 158P1D7 variant 4 is synthesized, coupled to KLH and used as immunogen. In another embodiment, peptides or bacterial fusion proteins are made that encompass the unique sequence generated by alternative splicing in the variants. In one example, a peptide encoding a consecutive sequence containing amino acids 682 and 683 in 158P1D7 variant 3 is used, such as amino acids 673-693. In another example, a peptide encoding a consecutive sequence containing amino acids 379-381 in 158P1D7 variant 6 is used, such as amino acids 369-391. Hybridomas are then selected that recognize the respective variant specific antigen and also recognize the full length variant protein expressed in cells. Such selection utilizes immunoassays described above such as Western blotting, immunoprecipitation, and flow cytometry.

To generate 158P1D7 monoclonal antibodies the following protocols were used. 5 Balb/c mice were immunized subcutaneously with 2µg of peptide in Quiagen ImmuneEasy™ adjuvant. Immunizations were given 2 weeks apart. The peptide used was a 12 amino acid peptide consisting of amino acids 274-285 with the sequence EEHEDPSGSLHL (SEQ ID NO: 41) conjugated to KLH at the C' terminal (Keyhole Limpet Hemocyanin).

B-cells from spleens of immunized mice were fused with the fusion partner Sp2/0 under the influence of polyethylene glycol. Antibody producing hybridomas were selected by screening on peptide coated ELISA plates indicating specific binding to the peptide and then by FACS on cells expressing 158P1D7. This produced and identified four 158P1D7 extra cellular domain (ECD) specific antibodies designated: M15-68(2)18.1.1; M15-68(2)22.1.1; M15-68(2)31.1.1 and M15-68(2)102.1.1.

The antibody designated M15-68(2)18.1.1 was sent (via Federal Express) to the American Type Culture Collection (ATCC), P.O. Box 1549, Manassas, VA 20108 on 06-February-2004 and assigned Accession number _____.
The characteristics of these four antibodies are set forth in Table LVII.

To clone the M15-68(2)18.1.1 antibody the following protocols were used. M15-68(2)18.1.1 hybridoma cells were lysed with Trizol reagent (Life Technologies, Gibco BRL). Total RNA was purified and quantified. First strand cDNAs was generated from total RNA with oligo (dT)12-18 priming using the Gibco-BRL Superscript Preamplification system. First strand cDNA was amplified using mouse Ig variable heavy chain primers, and mouse Ig variable light chain primers. PCR products were cloned into the pCRScript vector (Stratagene, La Jolla). Several clones were sequenced and the variable heavy (VH) and variable light (VL) chain regions determined. The nucleic acid and amino acid sequences of M15-68(2)18 variable heavy and light chain regions are set forth in Figure 34A and 34B and Figure 35A and 35B.

Example 10: HLA Class I and Class II Binding Assays

HLA class I and class II binding assays using purified HLA molecules are performed in accordance with disclosed protocols (*e.g.*, PCT publications WO 94/20127 and WO 94/03205; Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). Briefly, purified MHC molecules (5 to 500 nM) are incubated with various unlabeled peptide inhibitors and 1-10 nM ^{125}I-radiolabeled probe peptides as described. Following incubation, MHC-peptide complexes are separated from free peptide by gel filtration and the fraction of peptide bound is determined. Typically, in preliminary experiments, each MHC preparation is titered in the presence of

fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays are performed using these HLA concentrations.

Since under these conditions [label]<[HLA] and $IC_{50} \geq$[HLA], the measured IC_{50} values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 μg/ml to 1.2 ng/ml, and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC_{50} of a positive control for inhibition by the IC_{50} for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into IC_{50} nM values by dividing the IC_{50} nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation is accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

Binding assays as outlined above may be used to analyze HLA supermotif and/or HLA motif-bearing peptides.

Example 11: Identification of HLA Supermotif- and Motif-Bearing CTL Candidate Epitopes

HLA vaccine compositions of the invention can include multiple epitopes. The multiple epitopes can comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification and confirmation of supermotif- and motif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage is performed using the strategy described below.

Computer searches and algorithms for identification of supermotif and/or motif-bearing epitopes

The searches performed to identify the motif-bearing peptide sequences in the Example entitled "Antigenicity Profiles" and Tables V-XVIII and XXII-XLIX employ the protein sequence data from the gene product of 158P1D7 set forth in Figures 2 and 3.

Computer searches for epitopes bearing HLA Class I or Class II supermotifs or motifs are performed as follows. All translated 158P1D7 protein sequences are analyzed using a text string search software program to identify potential peptide sequences containing appropriate HLA binding motifs; such programs are readily produced in accordance with information in the art in view of known motif/supermotif disclosures. Furthermore, such calculations can be made mentally.

Identified A2-, A3-, and DR-supermotif sequences are scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms account for the impact of different amino acids at different positions, and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$\text{"}\Delta G\text{"} = a_{1i} \times a_{2i} \times a_{3i} \ldots\ldots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (*j*) at a given position (*i*) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue *j* occurs at position *i* in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide.

The method of derivation of specific algorithm coefficients has been described in Gulukota *et al., J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al., Human Immunol.* 45:79-93, 1996; and Southwood *et al., J. Immunol.* 160:3363-3373, 1998). Briefly, for all *i* positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying *j* is calculated relative to the remainder of the group, and used as the estimate of j_i. For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure.

To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

Selection of HLA-A2 supertype cross-reactive peptides

Complete protein sequences from 158P1D7 are scanned utilizing motif identification software, to identify 8-, 9- 10- and 11-mer sequences containing the HLA-A2-supermotif main anchor specificity. Typically, these sequences are then scored using the protocol described above and the peptides corresponding to the positive-scoring sequences are synthesized and tested for their capacity to bind purified HLA-A*0201 molecules *in vitro* (HLA-A*0201 is considered a prototype A2 supertype molecule).

These peptides are then tested for the capacity to bind to additional A2-supertype molecules (A*0202, A*0203, A*0206, and A*6802). Peptides that bind to at least three of the five A2-supertype alleles tested are typically deemed A2-supertype cross-reactive binders. Preferred peptides bind at an affinity equal to or less than 500 nM to three or more HLA-A2 supertype molecules.

Selection of HLA-A3 supermotif-bearing epitopes

The 158P1D7 protein sequence scanned above is also examined for the presence of peptides with the HLA-A3-supermotif primary anchors. Peptides corresponding to the HLA A3 supermotif-bearing sequences are then synthesized and tested for binding to HLA-A*0301 and HLA-A*1101 molecules, the molecules encoded by the two most prevalent A3-supertype alleles. The peptides that bind at least one of the two alleles with binding affinities of ≤500 nM, often ≤ 200 nM, are then tested for binding cross-reactivity to the other common A3-supertype alleles (e.g., A*3101, A*3301, and A*6801) to identify those that can bind at least three of the five HLA-A3-supertype molecules tested.

Selection of HLA-B7 supermotif bearing epitopes

The 158P1D7 protein is also analyzed for the presence of 8-, 9- 10-, or 11-mer peptides with the HLA-B7-supermotif. Corresponding peptides are synthesized and tested for binding to HLA-B*0702, the molecule encoded by the most common B7-supertype allele (*i.e.*, the prototype B7 supertype allele). Peptides binding B*0702 with IC_{50} of ≤500 nM are identified using standard methods. These peptides are then tested for binding to other common B7-supertype molecules (e.g., B*3501, B*5101, B*5301, and B*5401). Peptides capable of binding to three or more of the five B7-supertype alleles tested are thereby identified.

Selection of A1 and A24 motif-bearing epitopes

To further increase population coverage, HLA-A1 and -A24 epitopes can also be incorporated into vaccine compositions. An analysis of the 158P1D7 protein can also be performed to identify HLA-A1- and A24-motif-containing sequences.

High affinity and/or cross-reactive binding epitopes that bear other motif and/or supermotifs are identified using analogous methodology.

Example 12: Confirmation of Immunogenicity

Cross-reactive candidate CTL A2-supermotif-bearing peptides that are identified as described herein are selected to confirm *in vitro* immunogenicity. Confirmation is performed using the following methodology:

Target Cell Lines for Cellular Screening:

The .221A2.1 cell line, produced by transferring the HLA-A2.1 gene into the HLA-A, -B, -C null mutant human B-lymphoblastoid cell line 721.221, is used as the peptide-loaded target to measure activity of HLA-A2.1-restricted CTL. This cell line is grown in RPMI-1640 medium supplemented with antibiotics, sodium pyruvate, nonessential amino acids and 10%

(v/v) heat inactivated FCS. Cells that express an antigen of interest, or transfectants comprising the gene encoding the antigen of interest, can be used as target cells to confirm the ability of peptide-specific CTLs to recognize endogenous antigen.

Primary CTL Induction Cultures:

Generation of Dendritic Cells (DC): PBMCs are thawed in RPMI with 30 µg/ml DNAse, washed twice and resuspended in complete medium (RPMI-1640 plus 5% AB human serum, non-essential amino acids, sodium pyruvate, L-glutamine and penicillin/streptomycin). The monocytes are purified by plating 10 x 10^6 PBMC/well in a 6-well plate. After 2 hours at 37°C, the non-adherent cells are removed by gently shaking the plates and aspirating the supernatants. The wells are washed a total of three times with 3 ml RPMI to remove most of the non-adherent and loosely adherent cells. Three ml of complete medium containing 50 ng/ml of GM-CSF and 1,000 U/ml of IL-4 are then added to each well. TNFα is added to the DCs on day 6 at 75 ng/ml and the cells are used for CTL induction cultures on day 7.

Induction of CTL with DC and Peptide: CD8+ T-cells are isolated by positive selection with Dynal immunomagnetic beads (Dynabeads® M-450) and the detacha-bead® reagent. Typically about 200-250x10^6 PBMC are processed to obtain 24x10^6 $CD8^+$ T-cells (enough for a 48-well plate culture). Briefly, the PBMCs are thawed in RPMI with 30µg/ml DNAse, washed once with PBS containing 1% human AB serum and resuspended in PBS/1% AB serum at a concentration of 20x10^6cells/ml. The magnetic beads are washed 3 times with PBS/AB serum, added to the cells (140µl beads/20x10^6 cells) and incubated for 1 hour at 4°C with continuous mixing. The beads and cells are washed 4x with PBS/AB serum to remove the nonadherent cells and resuspended at 100x10^6 cells/ml (based on the original cell number) in PBS/AB serum containing 100µl/ml detacha-bead® reagent and 30 µg/ml DNAse. The mixture is incubated for 1 hour at room temperature with continuous mixing. The beads are washed again with PBS/AB/DNAse to collect the CD8+ T-cells. The DC are collected and centrifuged at 1300 rpm for 5-7 minutes, washed once with PBS with 1% BSA, counted and pulsed with 40µg/ml of peptide at a cell concentration of 1-2x10^6/ml in the presence of 3µg/ml $ß_2$- microglobulin for 4 hours at 20°C. The DC are then irradiated (4,200 rads), washed 1 time with medium and counted again.

Setting up induction cultures: 0.25 ml cytokine-generated DC (at 1x10^5 cells/ml) are co-cultured with 0.25ml of CD8+ T-cells (at 2x10^6 cell/ml) in each well of a 48-well plate in the presence of 10 ng/ml of IL-7. Recombinant human IL-10 is added the next day at a final concentration of 10 ng/ml and rhuman IL-2 is added 48 hours later at 10 IU/ml.

Restimulation of the induction cultures with peptide-pulsed adherent cells: Seven and fourteen days after the primary induction, the cells are restimulated with peptide-pulsed adherent cells. The PBMCs are thawed and washed twice with RPMI and DNAse. The cells are resuspended at 5x10^6 cells/ml and irradiated at ~4200 rads. The PBMCs are plated at 2x10^6 in 0.5 ml complete medium per well and incubated for 2 hours at 37°C. The plates are washed twice with RPMI by tapping the plate gently to remove the nonadherent cells and the adherent cells pulsed with 10µg/ml of peptide in the presence of 3 µg/ml $ß_2$ microglobulin in 0.25ml RPMI/5%AB per well for 2 hours at 37°C. Peptide solution from each well is aspirated and the wells are washed once with RPMI. Most of the media is aspirated from the induction cultures (CD8+ cells) and brought to 0.5 ml with fresh media. The cells are then transferred to the wells containing the peptide-pulsed adherent cells. Twenty four hours later recombinant human IL-10 is added at a final concentration of 10 ng/ml and recombinant human IL2 is added the next day and again 2-3 days later at 50IU/ml (Tsai *et al.*, *Critical Reviews in Immunology* 18(1-2):65-75, 1998). Seven days later, the cultures are assayed for CTL activity in a ^{51}Cr release assay. In some experiments the cultures are assayed for peptide-specific recognition in the *in situ* IFNγ ELISA at the time of the second restimulation followed by assay of endogenous recognition 7 days later. After expansion, activity is measured in both assays for a side-by-side comparison.

Measurement of CTL lytic activity by ^{51}Cr release.

Seven days after the second restimulation, cytotoxicity is determined in a standard (5 hr) ^{51}Cr release assay by assaying individual wells at a single E:T. Peptide-pulsed targets are prepared by incubating the cells with 10μg/ml peptide overnight at 37°C.

Adherent target cells are removed from culture flasks with trypsin-EDTA. Target cells are labeled with 200μCi of ^{51}Cr sodium chromate (Dupont, Wilmington, DE) for 1 hour at 37°C. Labeled target cells are resuspended at 10^6 per ml and diluted 1:10 with K562 cells at a concentration of $3.3x10^6$/ml (an NK-sensitive erythroblastoma cell line used to reduce non-specific lysis). Target cells (100 μl) and effectors (100μl) are plated in 96 well round-bottom plates and incubated for 5 hours at 37°C. At that time, 100 μl of supernatant are collected from each well and percent lysis is determined according to the formula:

[(cpm of the test sample- cpm of the spontaneous ^{51}Cr release sample)/(cpm of the maximal ^{51}Cr release sample- cpm of the spontaneous ^{51}Cr release sample)] x 100.

Maximum and spontaneous release are determined by incubating the labeled targets with 1% Trition X-100 and media alone, respectively. A positive culture is defined as one in which the specific lysis (sample- background) is 10% or higher in the case of individual wells and is 15% or more at the two highest E:T ratios when expanded cultures are assayed.

In situ Measurement of Human IFNγ Production as an Indicator of Peptide-specific and Endogenous Recognition

Immulon 2 plates are coated with mouse anti-human IFNγ monoclonal antibody (4 μg/ml 0.1M $NaHCO_3$, pH8.2) overnight at 4°C. The plates are washed with Ca^{2+}; Mg^{2+}-free PBS/0.05% Tween 20 and blocked with PBS/10% FCS for two hours, after which the CTLs (100 μl/well) and targets (100 μl/well) are added to each well, leaving empty wells for the standards and blanks (which received media only). The target cells, either peptide-pulsed or endogenous targets, are used at a concentration of $1x10^6$ cells/ml. The plates are incubated for 48 hours at 37°C with 5% CO_2.

Recombinant human IFN-gamma is added to the standard wells starting at 400 pg or 1200pg/100 microliter/well and the plate incubated for two hours at 37°C. The plates are washed and 100 μl of biotinylated mouse anti-human IFN-gamma monoclonal antibody (2 microgram/ml in PBS/3%FCS/0.05% Tween 20) are added and incubated for 2 hours at room temperature. After washing again, 100 microliter HRP-streptavidin (1:4000) are added and the plates incubated for one hour at room temperature. The plates are then washed 6x with wash buffer, 100 microliter/well developing solution (TMB 1:1) are added, and the plates allowed to develop for 5-15 minutes. The reaction is stopped with 50 microliter/well 1M H_3PO_4 and read at OD450. A culture is considered positive if it measured at least 50 pg of IFN-gamma/well above background and is twice the background level of expression.

CTL Expansion.

Those cultures that demonstrate specific lytic activity against peptide-pulsed targets and/or tumor targets are expanded over a two week period with anti-CD3. Briefly, $5x10^4$ CD8+ cells are added to a T25 flask containing the following: $1x10^6$ irradiated (4,200 rad) PBMC (autologous or allogeneic) per ml, $2x10^5$ irradiated (8,000 rad) EBV- transformed cells per ml, and OKT3 (anti-CD3) at 30ng per ml in RPMI-1640 containing 10% (v/v) human AB serum, non-essential amino acids, sodium pyruvate, 25μM 2-mercaptoethanol, L-glutamine and penicillin/streptomycin. Recombinant human IL2 is added 24 hours later at a final concentration of 200IU/ml and every three days thereafter with fresh media at 50IU/ml. The cells are split if the cell concentration exceeds $1x10^6$/ml and the cultures are assayed between days 13 and 15 at E:T ratios of 30, 10, 3 and 1:1 in the ^{51}Cr release assay or at $1x10^6$/ml in the *in situ* IFNγ assay using the same targets as before the expansion.

Cultures are expanded in the absence of anti-$CD3^+$ as follows. Those cultures that demonstrate specific lytic activity against peptide and endogenous targets are selected and $5x10^4$ $CD8^+$ cells are added to a T25 flask containing the following: $1x10^6$ autologous PBMC per ml which have been peptide-pulsed with 10 μg/ml peptide for two hours at 37°C and

irradiated (4,200 rad); $2x10^5$ irradiated (8,000 rad) EBV-transformed cells per ml RPMI-1640 containing 10%(v/v) human AB serum, non-essential AA, sodium pyruvate, 25mM 2-ME, L-glutamine and gentamicin.

Immunogenicity of A2 supermotif-bearing peptides

A2-supermotif cross-reactive binding peptides are tested in the cellular assay for the ability to induce peptide-specific CTL in normal individuals. In this analysis, a peptide is typically considered to be an epitope if it induces peptide-specific CTLs in at least individuals, and preferably, also recognizes the endogenously expressed peptide.

Immunogenicity can also be confirmed using PBMCs isolated from patients bearing a tumor that expresses 158P1D7. Briefly, PBMCs are isolated from patients, re-stimulated with peptide-pulsed monocytes and assayed for the ability to recognize peptide-pulsed target cells as well as transfected cells endogenously expressing the antigen.

Evaluation of A*03/A11 immunogenicity

HLA-A3 supermotif-bearing cross-reactive binding peptides are also evaluated for immunogenicity using methodology analogous for that used to evaluate the immunogenicity of the HLA-A2 supermotif peptides.

Evaluation of B7 immunogenicity

Immunogenicity screening of the B7-supertype cross-reactive binding peptides identified as set forth herein are confirmed in a manner analogous to the confirmation of A2-and A3-supermotif-bearing peptides.

Peptides bearing other supermotifs/motifs, *e.g.*, HLA-A1, HLA-A24 *etc.* are also confirmed using similar methodology

Example 13: Implementation of the Extended Supermotif to Improve the Binding Capacity of Native Epitopes by Creating Analogs

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in the identification and preparation of highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analoged to confer upon the peptide certain characteristics, *e.g.* greater cross-reactivity within the group of HLA molecules that comprise a supertype, and/or greater binding affinity for some or all of those HLA molecules. Examples of analoging peptides to exhibit modulated binding affinity are set forth in this example.

Analoging at Primary Anchor Residues

Peptide engineering strategies are implemented to further increase the cross-reactivity of the epitopes. For example, the main anchors of A2-supermotif-bearing peptides are altered, for example, to introduce a preferred L, I, V, or M at position 2, and I or V at the C-terminus.

To analyze the cross-reactivity of the analog peptides, each engineered analog is initially tested for binding to the prototype A2 supertype allele A*0201, then, if A*0201 binding capacity is maintained, for A2-supertype cross-reactivity.

Alternatively, a peptide is confirmed as binding one or all supertype members and then analogued to modulate binding affinity to any one (or more) of the supertype members to add population coverage.

The selection of analogs for immunogenicity in a cellular screening analysis is typically further restricted by the capacity of the parent wild type (WT) peptide to bind at least weakly, *i.e.*, bind at an IC_{50} of 5000nM or less, to three of more A2 supertype alleles. The rationale for this requirement is that the WT peptides must be present endogenously in sufficient quantity to be biologically relevant. Analoged peptides have been shown to have increased immunogenicity and cross-reactivity by T cells specific for the parent epitope (*see, e.g.*, Parkhurst *et al.*, *J. Immunol.* 157:2539, 1996; and Pogue *et al.*, *Proc. Natl. Acad. Sci. USA* 92:8166, 1995).

In the cellular screening of these peptide analogs, it is important to confirm that analog-specific CTLs are also able to recognize the wild-type peptide and, when possible, target cells that endogenously express the epitope.

Analoging of HLA-A3 and B7-supermotif-bearing peptides

Analogs of HLA-A3 supermotif-bearing epitopes are generated using strategies similar to those employed in analoging HLA-A2 supermotif-bearing peptides. For example, peptides binding to 3/5 of the A3-supertype molecules are engineered at primary anchor residues to possess a preferred residue (V, S, M, or A) at position 2.

The analog peptides are then tested for the ability to bind A*03 and A*11 (prototype A3 supertype alleles). Those peptides that demonstrate ≤ 500 nM binding capacity are then confirmed as having A3-supertype cross-reactivity.

Similarly to the A2- and A3- motif bearing peptides, peptides binding 3 or more B7-supertype alleles can be improved, where possible, to achieve increased cross-reactive binding or greater binding affinity or binding half life. B7 supermotif-bearing peptides are, for example, engineered to possess a preferred residue (V, I, L, or F) at the C-terminal primary anchor position, as demonstrated by Sidney *et al.* (*J. Immunol.* 157:3480-3490, 1996).

Analoging at primary anchor residues of other motif and/or supermotif-bearing epitopes is performed in a like manner.

The analog peptides are then be confirmed for immunogenicity, typically in a cellular screening assay. Again, it is generally important to demonstrate that analog-specific CTLs are also able to recognize the wild-type peptide and, when possible, targets that endogenously express the epitope.

Analoging at Secondary Anchor Residues

Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides and/or peptides that bind HLA molecules with increased affinity by identifying particular residues at secondary anchor positions that are associated with such properties. For example, the binding capacity of a B7 supermotif-bearing peptide with an F residue at position 1 is analyzed. The peptide is then analoged to, for example, substitute L for F at position 1. The analoged peptide is evaluated for increased binding affinity, binding half life and/or increased cross-reactivity. Such a procedure identifies analoged peptides with enhanced properties.

Engineered analogs with sufficiently improved binding capacity or cross-reactivity can also be tested for immunogenicity in HLA-B7-transgenic mice, following for example, IFA immunization or lipopeptide immunization. Analogued peptides are additionally tested for the ability to stimulate a recall response using PBMC from patients with 158P1D7-expressing tumors.

Other analoguing strategies

Another form of peptide analoguing, unrelated to anchor positions, involves the substitution of a cysteine with α-amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substitution of α-amino butyric acid for cysteine not only alleviates this problem, but has been shown to improve binding and crossbinding capabilities in some instances (*see, e.g.*, the review by Sette *et al.*, In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999).

Thus, by the use of single amino acid substitutions, the binding properties and/or cross-reactivity of peptide ligands for HLA supertype molecules can be modulated.

Example 14. Identification and confirmation of 158P1D7-derived sequences with HLA-DR binding motifs

Peptide epitopes bearing an HLA class II supermotif or motif are identified and confirmed as outlined below using methodology similar to that described for HLA Class I peptides.

Selection of HLA-DR-supermotif-bearing epitopes.

To identify 158P1D7-derived, HLA class II HTL epitopes, the 158P1D7 antigen is analyzed for the presence of sequences bearing an HLA-DR-motif or supermotif. Specifically, 15-mer sequences are selected comprising a DR-supermotif, comprising a 9-mer core, and three-residue N- and C-terminal flanking regions (15 amino acids total).

Protocols for predicting peptide binding to DR molecules have been developed (Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). These protocols, specific for individual DR molecules, allow the scoring, and ranking, of 9-mer core regions. Each protocol not only scores peptide sequences for the presence of DR-supermotif primary anchors (i.e., at position 1 and position 6) within a 9-mer core, but additionally evaluates sequences for the presence of secondary anchors. Using allele-specific selection tables (see, *e.g.*, Southwood *et al.*, *ibid.*), it has been found that these protocols efficiently select peptide sequences with a high probability of binding a particular DR molecule. Additionally, it has been found that performing these protocols in tandem, specifically those for DR1, DR4w4, and DR7, can efficiently select DR cross-reactive peptides.

The 158P1D7-derived peptides identified above are tested for their binding capacity for various common HLA-DR molecules. All peptides are initially tested for binding to the DR molecules in the primary panel: DR1, DR4w4, and DR7. Peptides binding at least two of these three DR molecules are then tested for binding to DR2w2 β1, DR2w2 β2, DR6w19, and DR9 molecules in secondary assays. Finally, peptides binding at least two of the four secondary panel DR molecules, and thus cumulatively at least four of seven different DR molecules, are screened for binding to DR4w15, DR5w11, and DR8w2 molecules in tertiary assays. Peptides binding at least seven of the ten DR molecules comprising the primary, secondary, and tertiary screening assays are considered cross-reactive DR binders. 158P1D7-derived peptides found to bind common HLA-DR alleles are of particular interest.

Selection of DR3 motif peptides

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is a relevant criterion in the selection of HTL epitopes. Thus, peptides shown to be candidates may also be assayed for their DR3 binding capacity. However, in view of the binding specificity of the DR3 motif, peptides binding only to DR3 can also be considered as candidates for inclusion in a vaccine formulation.

To efficiently identify peptides that bind DR3, target 158P1D7 antigens are analyzed for sequences carrying one of the two DR3-specific binding motifs reported by Geluk *et al.* (*J. Immunol.* 152:5742-5748, 1994). The corresponding peptides are then synthesized and confirmed as having the ability to bind DR3 with an affinity of 1μM or better, i.e., less than 1 μM. Peptides are found that meet this binding criterion and qualify as HLA class II high affinity binders.

DR3 binding epitopes identified in this manner are included in vaccine compositions with DR supermotif-bearing peptide epitopes.

Similarly to the case of HLA class I motif-bearing peptides, the class II motif-bearing peptides are analoged to improve affinity or cross-reactivity. For example, aspartic acid at position 4 of the 9-mer core sequence is an optimal residue for DR3 binding, and substitution for that residue often improves DR 3 binding.

Example 15: Immunogenicity of 158P1D7-derived HTL epitopes

This example determines immunogenic DR supermotif- and DR3 motif-bearing epitopes among those identified using the methodology set forth herein.

Immunogenicity of HTL epitopes are confirmed in a manner analogous to the determination of immunogenicity of CTL epitopes, by assessing the ability to stimulate HTL responses and/or by using appropriate transgenic mouse models. Immunogenicity is determined by screening for: 1.) *in vitro* primary induction using normal PBMC or 2.) recall responses from patients who have 158P1D7-expressing tumors.

Example 16: Calculation of phenotypic frequencies of HLA-supertypes in various ethnic backgrounds to determine breadth of population coverage

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles are determined. Gene frequencies for each HLA allele are calculated from antigen or allele frequencies utilizing the binomial distribution formulae gf=1-(SQRT(1-af)) (see, *e.g.*, Sidney *et al.*, *Human Immunol.* 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies are calculated, and the cumulative antigen frequencies derived by the use of the inverse formula [af=1-(1-Cgf)^2].

Where frequency data is not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies is assumed. To obtain total potential supertype population coverage no linkage disequilibrium is assumed, and only alleles confirmed to belong to each of the supertypes are included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations are made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (e.g., total=A+B*(1-A)). Confirmed members of the A3-like supertype are A3, A11, A31, A*3301, and A*6801. Although the A3-like supertype may also include A34, A66, and A*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed members of the A2-like supertype family are A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, and A*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B*3501-03, B51, B*5301, B*5401, B*5501-2, B*5601, B*6701, and B*7801 (potentially also B*1401, B*3504-06, B*4201, and B*5602).

Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups. Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%. An analogous approach can be used to estimate population coverage achieved with combinations of class II motif-bearing epitopes.

Immunogenicity studies in humans (*e.g.*, Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997; Doolan *et al.*, *Immunity* 7:97, 1997; and Threlkeld *et al.*, *J. Immunol.* 159:1648, 1997) have shown that highly cross-reactive binding peptides are almost always recognized as epitopes. The use of highly cross-reactive binding peptides is an important selection criterion in identifying candidate epitopes for inclusion in a vaccine that is immunogenic in a diverse population.

With a sufficient number of epitopes (as disclosed herein and from the art), an average population coverage is predicted to be greater than 95% in each of five major ethnic populations. The game theory Monte Carlo simulation analysis, which is known in the art (see *e.g.*, Osborne, M.J. and Rubinstein, A. "A course in game theory" MIT Press, 1994), can be used to estimate what percentage of the individuals in a population comprised of the Caucasian, North American Black, Japanese, Chinese, and Hispanic ethnic groups would recognize the vaccine epitopes described herein. A preferred percentage is 90%. A more preferred percentage is 95%.

Example 17: CTL Recognition Of Endogenously Processed Antigens After Priming

This example confirms that CTL induced by native or analoged peptide epitopes identified and selected as described herein recognize endogenously synthesized, *i.e.*, native antigens.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes, for example HLA-A2 supermotif-bearing epitopes, are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are

assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ^{51}Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ^{51}Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with 158P1D7 expression vectors.

The results demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized 158P1D7 antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that are being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

Example 18: Activity of CTL-HTL Conjugated Epitopes In Transgenic Mice

This example illustrates the induction of CTLs and HTLs in transgenic mice, by use of a 158P1D7-derived CTL and HTL peptide vaccine compositions. The vaccine composition used herein comprise peptides to be administered to a patient with a 158P1D7-expressing tumor. The peptide composition can comprise multiple CTL and/or HTL epitopes. The epitopes are identified using methodology as described herein. This example also illustrates that enhanced immunogenicity can be achieved by inclusion of one or more HTL epitopes in a CTL vaccine composition; such a peptide composition can comprise an HTL epitope conjugated to a CTL epitope. The CTL epitope can be one that binds to multiple HLA family members at an affinity of 500 nM or less, or analogs of that epitope. The peptides may be lipidated, if desired.

Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are used to confirm the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, and are primed subcutaneously (base of the tail) with a 0.1 ml of peptide in Incomplete Freund's Adjuvant, or if the peptide composition is a lipidated CTL/HTL conjugate, in DMSO/saline, or if the peptide composition is a polypeptide, in PBS or Incomplete Freund's Adjuvant. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (*e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991)

In vitro CTL activation: One week after priming, spleen cells ($30x10^6$ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts ($10x10^6$ cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

Assay for cytotoxic activity: Target cells (1.0 to $1.5x10^6$) are incubated at 37°C in the presence of 200 µl of ^{51}Cr. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1 µg/ml. For the assay, 10^4 ^{51}Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 µl) in U-bottom 96-well plates. After a six hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release = 100 x (experimental release - spontaneous release)/(maximum release - spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ^{51}Cr release data is expressed as lytic units/10^6 cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a six hour ^{51}Cr release assay. To obtain specific lytic units/10^6, the lytic units/10^6 obtained in the absence of peptide is subtracted from the lytic units/10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., $5x10^5$ effector cells for 10,000

targets) in the absence of peptide and 5:1 (i.e., $5x10^4$ effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000)-(1/500,000)] \times 10^6 = 18$ LU.

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using, for example, CTL epitopes as outlined above in the Example entitled "Confirmation of Immunogenicity". Analyses similar to this may be performed to confirm the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures, it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

Example 19: Selection of CTL and HTL epitopes for inclusion in an 158P1D7-specific vaccine.

This example illustrates a procedure for selecting peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (*i.e.*, minigene) that encodes peptide(s), or can be single and/or polyepitopic peptides.

The following principles are utilized when selecting a plurality of epitopes for inclusion in a vaccine composition. Each of the following principles is balanced in order to make the selection.

Epitopes are selected which, upon administration, mimic immune responses that are correlated with 158P1D7 clearance. The number of epitopes used depends on observations of patients who spontaneously clear 158P1D7. For example, if it has been observed that patients who spontaneously clear 158P1D7 generate an immune response to at least three (3) from 158P1D7 antigen, then three or four (3-4) epitopes should be included for HLA class I. A similar rationale is used to determine HLA class II epitopes.

Epitopes are often selected that have a binding affinity of an IC_{50} of 500 nM or less for an HLA class I molecule, or for class II, an IC_{50} of 1000 nM or less; or HLA Class I peptides with high binding scores from the BIMAS web site, at URL bimas.dcrt.nih.gov/.

In order to achieve broad coverage of the vaccine through out a diverse population, sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. In one embodiment, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

When creating polyepitopic compositions, or a minigene that encodes same, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes. For example, a protein sequence for the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. Epitopes may be nested or overlapping (*i.e.*, frame shifted relative to one another). For example, with overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Each epitope can be exposed and bound by an HLA molecule upon administration of such a peptide. A multi-epitopic, peptide can be generated synthetically, recombinantly, or via cleavage from the native source. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes. This embodiment provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent the creating of any analogs) directs the immune response to multiple peptide sequences that are actually present in 158P1D7, thus avoiding the

need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions. Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude to an immune response that controls or clears cells that bear or overexpress 158P1D7.

Example 20: Construction of "Minigene" Multi-Epitope DNA Plasmids

This example discusses the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of B cell, CTL and/or HTL epitopes or epitope analogs as described herein.

A minigene expression plasmid typically includes multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes. HLA class I supermotif or motif-bearing peptide epitopes derived 158P1D7, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from 158P1D7 to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

Such a construct may additionally include sequences that direct the HTL epitopes to the endoplasmic reticulum. For example, the Ii protein may be fused to one or more HTL epitopes as described in the art, wherein the CLIP sequence of the Ii protein is removed and replaced with an HLA class II epitope sequence so that HLA class II epitope is directed to the endoplasmic reticulum, where the epitope binds to an HLA class II molecules.

This example illustrates the methods to be used for construction of a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid of this example contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

Overlapping oligonucleotides that can, for example, average about 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated Tm of each primer pair) for 30 sec, and 72°C for 1 min.

For example, a minigene is prepared as follows. For a first PCR reaction, 5 μg of each of two oligonucleotides are annealed and extended: In an example using eight oligonucleotides, i.e., four pairs of primers, oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 μl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM $(NH4)_2SO_4$, 20 mM Tris-chloride, pH 8.75, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 μg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

Example 21: The Plasmid Construct and the Degree to Which It Induces Immunogenicity.

The degree to which a plasmid construct, for example a plasmid constructed in accordance with the previous Example, is able to induce immunogenicity is confirmed *in vitro* by determining epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (*see, e.g.*, Sijts *et al.*, *J. Immunol.* 156:683-692, 1996; Demotz *et al.*, *Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by diseased or transfected target cells, and then determining the concentration of peptide necessary to obtain equivalent levels of lysis or lymphokine release (*see, e.g.*, Kageyama *et al.*, *J. Immunol.* 154:567-576, 1995).

Alternatively, immunogenicity is confirmed through *in vivo* injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which are analyzed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in Alexander *et al.*, *Immunity* 1:751-761, 1994.

For example, to confirm the capacity of a DNA minigene construct containing at least one HLA-A2 supermotif peptide to induce CTLs *in vivo*, HLA-A2.1/K^b transgenic mice, for example, are immunized intramuscularly with 100 μg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ^{51}Cr release assay. The results indicate the magnitude of the CTL response directed against the A2-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and polyepitopic vaccine.

It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A2 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes, whereby it is also found that the minigene elicits appropriate immune responses directed toward the provided epitopes.

To confirm the capacity of a class II epitope-encoding minigene to induce HTLs *in vivo*, DR transgenic mice, or for those epitopes that cross react with the appropriate mouse MHC molecule, I-A^b-restricted mice, for example, are immunized intramuscularly with 100 μg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant. CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ^3H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander et al. Immunity 1:751-761, 1994). The results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

DNA minigenes, constructed as described in the previous Example, can also be confirmed as a vaccine in combination with a boosting agent using a prime boost protocol. The boosting agent can consist of recombinant protein (*e.g.*, Barnett *et al.*, *Aids Res. and Human Retroviruses* 14, Supplement 3:S299-S309, 1998) or recombinant vaccinia, for example, expressing a minigene or DNA encoding the complete protein of interest (*see, e.g.*, Hanke *et al.*, *Vaccine* 16:439-

445, 1998; Sedegah *et al.*, *Proc. Natl. Acad. Sci USA* 95:7648-53, 1998; Hanke and McMichael, *Immunol. Letters* 66:177-181, 1999; and Robinson *et al.*, *Nature Med.* 5:526-34, 1999).

For example, the efficacy of the DNA minigene used in a prime boost protocol is initially evaluated in transgenic mice. In this example, A2.1/K^b transgenic mice are immunized IM with 100 μg of a DNA minigene encoding the immunogenic peptides including at least one HLA-A2 supermotif-bearing peptide. After an incubation period (ranging from 3-9 weeks), the mice are boosted IP with 10^7 pfu/mouse of a recombinant vaccinia virus expressing the same sequence encoded by the DNA minigene. Control mice are immunized with 100 μg of DNA or recombinant vaccinia without the minigene sequence, or with DNA encoding the minigene, but without the vaccinia boost. After an additional incubation period of two weeks, splenocytes from the mice are immediately assayed for peptide-specific activity in an ELISPOT assay. Additionally, splenocytes are stimulated *in vitro* with the A2-restricted peptide epitopes encoded in the minigene and recombinant vaccinia, then assayed for peptide-specific activity in an alpha, beta and/or gamma IFN ELISA.

It is found that the minigene utilized in a prime-boost protocol elicits greater immune responses toward the HLA-A2 supermotif peptides than with DNA alone. Such an analysis can also be performed using HLA-A11 or HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 or HLA-B7 motif or supermotif epitopes. The use of prime boost protocols in humans is described below in the Example entitled "Induction of CTL Responses Using a Prime Boost Protocol."

Example 22: Peptide Composition for Prophylactic Uses

Vaccine compositions of the present invention can be used to prevent 158P1D7 expression in persons who are at risk for tumors that bear this antigen. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes such as those selected in the above Examples, which are also selected to target greater than 80% of the population, is administered to individuals at risk for a 158P1D7-associated tumor.

For example, a peptide-based composition is provided as a single polypeptide that encompasses multiple epitopes. The vaccine is typically administered in a physiological solution that comprises an adjuvant, such as Incomplete Freunds Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 μg, generally 100-5,000 μg, for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against 158P1D7-associated disease.

Alternatively, a composition typically comprising transfecting agents is used for the administration of a nucleic acid-based vaccine in accordance with methodologies known in the art and disclosed herein.

Example 23: Polyepitopic Vaccine Compositions Derived from Native 158P1D7 Sequences

A native 158P1D7 polyprotein sequence is analyzed, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes. The "relatively short" regions are preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct or overlapping, "nested" epitopes is selected; it can be used to generate a minigene construct. The construct is engineered to express the peptide, which corresponds to the native protein sequence. The "relatively short" peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping (*i.e.*, frame shifted relative to one

another). For example, with overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

The vaccine composition will include, for example, multiple CTL epitopes from 158P1D7 antigen and at least one HTL epitope. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (excluding an analoged embodiment) directs the immune response to multiple peptide sequences that are actually present in native 158P1D7, thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing peptide or nucleic acid vaccine compositions.

Related to this embodiment, computer programs are available in the art which can be used to identify in a target sequence, the greatest number of epitopes per sequence length.

Example 24: Polyepitopic Vaccine Compositions From Multiple Antigens

The 158P1D7 peptide epitopes of the present invention are used in conjunction with epitopes from other target tumor-associated antigens, to create a vaccine composition that is useful for the prevention or treatment of cancer that expresses 158P1D7 and such other antigens. For example, a vaccine composition can be provided as a single polypeptide that incorporates multiple epitopes from 158P1D7 as well as tumor-associated antigens that are often expressed with a target cancer associated with 158P1D7 expression, or can be administered as a composition comprising a cocktail of one or more discrete epitopes. Alternatively, the vaccine can be administered as a minigene construct or as dendritic cells which have been loaded with the peptide epitopes *in vitro*.

Example 25: Use of peptides to evaluate an immune response

Peptides of the invention may be used to analyze an immune response for the presence of specific antibodies, CTL or HTL directed to 158P1D7. Such an analysis can be performed in a manner described by Ogg *et al., Science* 279:2103-2106, 1998. In this Example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

In this example highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, 158P1D7 HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of disease or following immunization comprising an 158P1D7 peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey *et al., N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5' triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300g for 5 minutes and resuspended in 50 μl of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive non-diseased donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the 158P1D7 epitope, and thus the status of exposure to 158P1D7, or exposure to a vaccine that elicits a protective or therapeutic response.

Example 26: Use of Peptide Epitopes to Evaluate Recall Responses

The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from 158P1D7-associated disease or who have been vaccinated with an 158P1D7 vaccine.

For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any 158P1D7 vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 μg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 μg/ml to each well and HBV core 128-140 epitope is added at 1 μg/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format, 4 x 10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 μl/well of complete RPMI. On days 3 and 10, 100 ul of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with non-diseased control subjects as previously described (Rehermann, *et al., Nature Med.* 2:1104,1108, 1996; Rehermann *et al., J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al. J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al. J. Virol.* 66:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 μM, and labeled with 100 μCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: 100 x [(experimental release-spontaneous release)/maximum release-

spontaneous release)]. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to 158P1D7 or an 158P1D7 vaccine.

Similarly, Class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of $1.5x10^5$ cells/well and are stimulated with 10 μg/ml synthetic peptide of the invention, whole 158P1D7 antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 μCi ^{3}H-thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^{3}H-thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^{3}H-thymidine incorporation in the presence of antigen divided by the ^{3}H-thymidine incorporation in the absence of antigen.

Example 27: Induction Of Specific CTL Response In Humans

A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

A total of about 27 individuals are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

The vaccine is found to be both safe and efficacious.

Example 28: Phase II Trials In Patients Expressing 158P1D7

Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to patients having cancer that expresses 158P1D7. The main objectives of the trial are to determine an effective dose and regimen for inducing CTLs in cancer patients that express 158P1D7, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of these patients, as manifested, e.g., by the reduction and/or shrinking of lesions. Such a study is designed, for example, as follows:

The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot

of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65 and represent diverse ethnic backgrounds. All of them have a tumor that expresses 158P1D7.

Clinical manifestations or antigen-specific T-cell responses are monitored to assess the effects of administering the peptide compositions. The vaccine composition is found to be both safe and efficacious in the treatment of 158P1D7-associated disease.

Example 29: Induction of CTL Responses Using a Prime Boost Protocol

A prime boost protocol similar in its underlying principle to that used to confirm the efficacy of a DNA vaccine in transgenic mice, such as described above in the Example entitled "The Plasmid Construct and the Degree to Which It Induces Immunogenicity," can also be used for the administration of the vaccine to humans. Such a vaccine regimen can include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

For example, the initial immunization may be performed using an expression vector, such as that constructed in the Example entitled "Construction of 'Minigene' Multi-Epitope DNA Plasmids" in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster can be recombinant fowlpox virus administered at a dose of 5-10^7 to 5x10^9 pfu. An alternative recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples are obtained before immunization as well as at intervals following administration of the initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

Analysis of the results indicates that a magnitude of response sufficient to achieve a therapeutic or protective immunity against 158P1D7 is generated.

Example 30: Administration of Vaccine Compositions Using Dendritic Cells (DC)

Vaccines comprising peptide epitopes of the invention can be administered using APCs, or "professional" APCs such as DC. In this example, peptide-pulsed DC are administered to a patient to stimulate a CTL response *in vivo*. In this method, dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy or facilitate destruction, respectively, of the target cells that bear the 158P1D7 protein from which the epitopes in the vaccine are derived.

For example, a cocktail of epitope-comprising peptides is administered *ex vivo* to PBMC, or isolated DC therefrom. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides, and prior to reinfusion into patients, the DC are washed to remove unbound peptides.

As appreciated clinically, and readily determined by one of skill based on clinical outcomes, the number of DC reinfused into the patient can vary (see, *e.g.*, *Nature Med.* 4:328, 1998; *Nature Med.* 2:52, 1996 and *Prostate* 32:272, 1997). Although 2-50 x 10^6 DC per patient are typically administered, larger number of DC, such as 10^7 or 10^8 can also be provided. Such cell populations typically contain between 50-90% DC.

In some embodiments, peptide-loaded PBMC are injected into patients without purification of the DC. For example, PBMC generated after treatment with an agent such as Progenipoietin™ are injected into patients without purification of the DC. The total number of PBMC that are administered often ranges from 10^8 to 10^{10}. Generally, the cell doses injected into patients is based on the percentage of DC in the blood of each patient, as determined, for example, by immunofluorescence analysis with specific anti-DC antibodies. Thus, for example, if Progenipoietin™ mobilizes 2% DC in the peripheral blood of a given patient, and that patient is to receive 5 x 10^6 DC, then the patient will be injected with a total of 2.5 x 10^8 peptide-loaded PBMC. The percent DC mobilized by an agent such as Progenipoietin™ is typically estimated to be between 2-10%, but can vary as appreciated by one of skill in the art.

Ex vivo activation of CTL/HTL responses

Alternatively, *ex vivo* CTL or HTL responses to 158P1D7 antigens can be induced by incubating, in tissue culture, the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of APC, such as DC, and immunogenic peptides. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cells, *i.e.*, tumor cells.

Example 31: An Alternative Method of Identifying and Confirming Motif-Bearing Peptides

Another method of identifying and confirming motif-bearing peptides is to elute them from cells bearing defined MHC molecules. For example, EBV transformed B cell lines used for tissue typing have been extensively characterized to determine which HLA molecules they express. In certain cases these cells express only a single type of HLA molecule. These cells can be transfected with nucleic acids that express the antigen of interest, *e.g.* 158P1D7. Peptides produced by endogenous antigen processing of peptides produced as a result of transfection will then bind to HLA molecules within the cell and be transported and displayed on the cell's surface. Peptides are then eluted from the HLA molecules by exposure to mild acid conditions and their amino acid sequence determined, *e.g.*, by mass spectral analysis (*e.g.*, Kubo *et al.*, *J. Immunol.* 152:3913, 1994). Because the majority of peptides that bind a particular HLA molecule are motif-bearing, this is an alternative modality for obtaining the motif-bearing peptides correlated with the particular HLA molecule expressed on the cell.

Alternatively, cell lines that do not express endogenous HLA molecules can be transfected with an expression construct encoding a single HLA allele. These cells can then be used as described, *i.e.*, they can then be transfected with nucleic acids that encode 158P1D7 to isolate peptides corresponding to 158P1D7 that have been presented on the cell surface. Peptides obtained from such an analysis will bear motif(s) that correspond to binding to the single HLA allele that is expressed in the cell.

As appreciated by one in the art, one can perform a similar analysis on a cell bearing more than one HLA allele and subsequently determine peptides specific for each HLA allele expressed. Moreover, one of skill would also recognize that means other than transfection, such as loading with a protein antigen, can be used to provide a source of antigen to the cell.

Example 32: Complementary Polynucleotides

Sequences complementary to the 158P1D7-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring 158P1D7. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using, e.g., OLIGO 4.06 software (National Biosciences) and the coding sequence of 158P1D7. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the 158P1D7-encoding transcript.

Example 33: Purification of Naturally-occurring or Recombinant 158P1D7 Using 158P1D7 Specific Antibodies

Naturally occurring or recombinant 158P1D7 is substantially purified by immunoaffinity chromatography using antibodies specific for 158P1D7. An immunoaffinity column is constructed by covalently coupling anti-158P1D7 antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing 158P1D7 are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of 158P1D7 (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/158P1D7 binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GCR.P is collected.

Example 34: Identification of Molecules Which Interact with 158P1D7

158P1D7, or biologically active fragments thereof, are labeled with 121 1 Bolton-Hunter reagent.

(See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled 158P1D7, washed, and any wells with labeled 158P1D7 complex are assayed. Data obtained using different concentrations of 158P1D7 are used to calculate values for the number, affinity, and association of 158P1D7 with the candidate molecules. Throughout this application, various website data content, publications, applications and patents are referenced. (Websites are referenced by their Uniform Resource Locator, or URL, addresses on the World Wide Web.) The disclosures of each of these items of information are hereby incorporated by reference herein in their entireties.

Example 35: *In Vivo* Assay for 158P1D7 Tumor Growth Promotion

The effect of the 158P1D7 protein on tumor cell growth can be confirmed *in vivo* by gene overexpression in bladder cancer cells. For example, SCID mice can be injected SQ on each flank with 1 x 10^6 bladder cancer cells (such as SCaBER, UM-UC-3, HT1376, RT4, T24, TCC-SUP, J82 and SW780 cells) containing tkNeo empty vector or 158P1D7.

At least two strategies may be used: (1) Constitutive 158P1D7 expression under regulation of a promoter such as a constitutive promoter obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), or from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, provided such promoters are compatible with the host cell systems. (2) Regulated expression under control of an inducible vector system, such as ecdysone, tet, etc., can be used provided such promoters are compatible with the host cell systems. Tumor volume is then monitored at the appearance of palpable tumors and is followed over time to determine if 158P1D7-expressing cells grow at a faster rate and whether tumors produced by

158P1D7-expressing cells demonstrate characteristics of altered aggressiveness (e.g. enhanced metastasis, vascularization, reduced responsiveness to chemotherapeutic drugs). Additionally, mice can be implanted with the same cells orthotopically to determine if 158P1D7 has an effect on local growth in the bladder or on the ability of the cells to metastasize, specifically to lungs or lymph nodes (Fu, X., *et al.*, Int. J. Cancer, 1991. **49**: p. 938-939; Chang, S., *et al.*, Anticancer Res., 1997. **17**: p. 3239-3242; Peralta, E. A., *et al.*, J. Urol., 1999. **162**: p. 1806-1811). Furthermore, this assay is useful to confirm the 158P1D7 inhibitory effect of candidate therapeutic compositions, such as for example, 158P1D7 antibodies or intrabodies, and 158P1D7 antisense molecules or ribozymes.

The assay was performed using the following protocols. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and maintained in a strictly controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals. 158P1D7 transfected UM-UC-3 cells and parental cells were injected into the subcutaneous space of SCID mice. Each mouse received 4 x 10^6 cells suspended in 50% (v/v) of Matrigel. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula $W^2 \times L/2$. The Mann-Whitney U test was used to evaluate differences of tumor growth. All tests were two sided with α=0.05. The results show that 158P1D7 enhances the growth of bladder cancer in mice (Figure 27).

Example 36: 158P1D7 Monoclonal Antibody-mediated Inhibition of Bladder and Prostate Tumors *In Vivo*

The significant expression of 158P1D7 in cancer tissues, together with its restricted expression in normal tissues, makes 158P1D7 an excellent target for antibody therapy. In cases where the monoclonal antibody target is a cell surface protein, antibodies have been shown to be efficacious at inhibiting tumor growth (See, e.g., (Saffran, D., *et al.*, PNAS 10:1073-1078 or URL: pnas.org/cgi/doi/10.1073/pnas.051624698). In cases where the target is not on the cell surface, such as PSA and PAP in prostate cancer, antibodies have still been shown to recognize and inhibit growth of cells expressing those proteins (Saffran, D.C., *et al.*, Cancer and Metastasis Reviews, 1999. 18: p. 437-449). As with any cellular protein with a restricted expression profile, 158P1D7 is a target for T cell-based immunotherapy.

Accordingly, the therapeutic efficacy of anti-158P1D7 mAbs in human bladder cancer mouse models is modeled in 158P1D7-expressing bladder cancer xenografts or bladder cancer cell lines, such as those described in Example (the Example entitled "*In Vivo* Assay for 158P1D7 Tumor Growth Promotion", that have been engineered to express 158P1D7.

Antibody efficacy on tumor growth and metastasis formation is confirmed, e.g., in a mouse orthotopic bladder cancer xenograft model. The antibodies can be unconjugated, as discussed in this Example, or can be conjugated to a therapeutic modality, as appreciated in the art. It is confirmed that anti-158P1D7 mAbs inhibit formation of 158P1D7-expressing bladder and prostate tumors (Figures 30 and 31). Anti-158P1D7 mAbs can be tested for the retardation of the growth of established orthotopic tumors and the prolonged survival of tumor-bearing mice. These results indicate the utility of anti-158P1D7 mAbs in the treatment of local and advanced stages of bladder and prostate cancers. (See, e.g., Saffran, D., et al., PNAS 10:1073-1078 or URL: pnas.org/cgi/doi/10.1073/pnas.051624698)

Administration of anti-158P1D7 mAbs retard established orthotopic tumor growth and inhibit metastasis to distant sites, resulting in a significant prolongation in the survival of tumor-bearing mice. These studies indicate that 158P1D7 is an attractive target for immunotherapy and demonstrate the therapeutic potential of anti-158P1D7 mAbs for the treatment of local and metastatic bladder cancer.

This example demonstrates that unconjugated 158P1D7 monoclonal antibodies effectively to inhibit the growth of human bladder tumors grown in SCID mice; accordingly a combination of such efficacious monoclonal antibodies is also effective.

Tumor inhibition using multiple unconjugated 158P1D7 mAbs

Materials and Methods

158P1D7 Monoclonal Antibodies:

Monoclonal antibodies are raised against 158P1D7 as described in the Example entitled "Generation of 158P1D7 Monoclonal Antibodies (mAbs)." The antibodies are characterized by ELISA, Western blot, FACS, and immunoprecipitation, in accordance with techniques known in the art, for their capacity to bind 158P1D7. Epitope mapping data for the anti-158P1D7 mAbs, as determined by ELISA and Western analysis, recognize epitopes on the 158P1D7 protein. Immunohistochemical analysis of bladder cancer tissues and cells with these antibodies is performed.

The monoclonal antibodies are purified from ascites or hybridoma tissue culture supernatants by Protein-G Sepharose chromatography, dialyzed against PBS, filter sterilized, and stored at -20°C. Protein determinations are performed by a Bradford assay (Bio-Rad, Hercules, CA). A therapeutic monoclonal antibody or a cocktail comprising a mixture of individual monoclonal antibodies is prepared and used for the treatment of mice receiving subcutaneous or orthotopic injections of bladder tumor xenografts.

Bladder Cancer Cell Lines

Bladder cancer cell lines (Scaber, J82, UM-UC-3, HT1376, RT4, T24, TCC-SUP, J82 and SW780) expressing 158P1D7 are generated by retroviral gene transfer as described in Hubert, R.S., et al., STEAP: a prostate-specific cell-surface antigen highly expressed in human prostate tumors. Proc Natl Acad Sci U S A, 1999. **96**(25):14523-8. Anti-158P1D7 staining is detected by using an FITC-conjugated goat anti-mouse antibody (Southern Biotechnology Associates) followed by analysis on a Coulter Epics-XL f low cytometer.

In Vivo Mouse Models.

Subcutaneous (s.c.) tumors are generated by injection of 1 x 10^6 158P1D7-expressing bladder cancer cells mixed at a 1:1 dilution with Matrigel (Collaborative Research) in the right flank of male SCID mice. To test antibody efficacy on tumor formation, i.p. antibody injections are started on the same day as tumor-cell injections. As a control, mice are injected with either purified mouse IgG (ICN) or PBS; or a purified monoclonal antibody that recognizes an irrelevant antigen not expressed in human cells. In preliminary studies, no difference is found between mouse IgG or PBS on tumor growth. Tumor sizes are determined by vernier caliper measurements, and the tumor volume is calculated as length x width x height. Mice with s.c. tumors greater than 1.5 cm in diameter are sacrificed. Circulating levels of anti-158P1D7 mAbs are determined by a capture ELISA kit (Bethyl Laboratories, Montgomery, TX). (See, e.g., (Saffran, D., et al., PNAS 10:1073-1078)

Orthotopic injections are performed, for example, in two alternative embodiments, under anesthesia by, for example, use of ketamine/xylazine. In a first embodiment, an intravesicular injection of bladder cancer cells is administered directly through the urethra and into the bladder (Peralta, E. A., *et al.*, J. Urol., 1999. 162:1806-1811). In a second embodiment, an incision is made through the abdominal wall, the bladder is exposed, and bladder tumor tissue pieces (1-2 mm in size) derived from a s.c. tumor are surgically glued onto the exterior wall of the bladder, termed "onplantation" (Fu, X., *et al.*, Int. J. Cancer, 1991. 49: 938-939; Chang, S., *et al.*, Anticancer Res., 1997. 17: p. 3239-3242). Antibodies can be administered to groups of mice at the time of tumor injection or onplantation, or after 1-2 weeks to allow tumor establishment.

Anti-158P1D7 mAbs Inhibit Growth of 158P1D7-Expressing Bladder Cancer Tumors

In one embodiment, the effect of anti-158P1D7 mAbs on tumor formation is tested by using the bladder onplantation orthotopic model. As compared with the s.c. tumor model, the orthotopic model, which requires surgical attachment of tumor tissue directly on the bladder, results in a local tumor growth, development of metastasis in distal sites, and subsequent death (Fu, X., *et al.*, Int. J. Cancer, 1991. 49: p. 938-939; Chang, S., *et al.*, Anticancer Res., 1997. 17: p.

3239-3242). This feature make the orthotopic model more representative of human disease progression and allows one to follow the therapeutic effect of mAbs, as well as other therapeutic modalities, on clinically relevant end points.

Accordingly, 158P1D7-expressing tumor cells are onplanted orthotopically, and 2 days later, the mice are segregated into two groups and treated with either: a) 50-2000μg, usually 200-500μg, of anti-158P1D7 Ab, or b) PBS, three times per week for two to five weeks. Mice are monitored weekly for indications of tumor growth.

As noted, a major advantage of the orthotopic bladder cancer model is the ability to study the development of metastases. Formation of metastasis in mice bearing established orthotopic tumors is studied by histological analysis of tissue sections, including lung and lymph nodes (Fu, X., *et al.*, Int. J. Cancer, 1991. 49:938-939; Chang, S., *et al.*, Anticancer Res., 1997. 17:3239-3242). Additionally, IHC analysis using anti-158P1D7 antibodies can be performed on the tissue sections.

Mice bearing established orthotopic 158P1D7-expressing bladder tumors are administered 1000μg injections of either anti-158P1D7 mAb or PBS over a 4-week period. Mice in both groups are allowed to establish a high tumor burden (1-2 weeks growth), to ensure a high frequency of metastasis formation in mouse lungs and lymph nodes. Mice are then sacrificed and their local bladder tumor and lung and lymph node tissue are analyzed for the presence of tumor cells by histology and IHC analysis.

In another embodiment, the effect of anti-158P1D7 mAbs on tumor growth was tested using the following protocols. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and were maintained in a strictly-controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals.

UG-B1, a patient bladder cancer, was used to establish xenograft models. Stock tumors regularly maintained in SCID mice were sterilely dissected, minced, and digested using Pronase (Calbiochem, San Diego, CA). Cell suspensions generated were incubated overnight at 37°C to obtain a homogeneous single-cell suspension. Each mouse received 2.5 x 10^6 cells at the subcutaneous site of right flank. Murine monoclonal antibodies to 158P1D7 were tested at a dose of 500 μg/mouse in the study. PBS was used as control. MAbs were dosed intra-peritoneally twice a week for a total of 12 doses, starting on the same day of tumor cell injection. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula: $W^2 \times L/2$. The results show that Anti-158P1D7 mAbs are capable of inhibiting the growth of human bladder carcinoma in mice (Figure 30).

Anti-158P1D7 mAbs retard the Growth of established 158P1D7-Expressing Prostate Cancer Tumors

In another embodiment, the effect of anti-158P1D7 mAbs on tumor growth was tested using the following protocols. Male ICR-SCID mice, 5-6 weeks old (Charles River Laboratory, Wilmington, MA) were used and were maintained in a strictly-controlled environment in accordance with the NIH Guide for the Care and Use of Laboratory Animals. LAPC-9AD, an androgen-dependent human prostate cancer, was used to establish xenograft models. Stock tumors were regularly maintained in SCID mice. At the day of implantation, stock tumors were harvested and trimmed of necrotic tissues and minced to 1 mm^3 pieces. Each mouse received 4 pieces of tissues at the subcutaneous site of right flank. Murine monoclonal antibodies to 158P1D7 were tested at a dose of 500 μg/mouse and 500 μg/mouse respectively. PBS and anti-KLH monoclonal antibody were used as controls. The study cohort consisted of 4 groups with 6 mice in each group. MAbs were dosed intra-peritoneally twice a week for a total of 8 doses. Treatment was started when tumor volume reached 45 mm^3. Tumor size was monitored through caliper measurements twice a week. The longest dimension (L) and the dimension perpendicular to it (W) were taken to calculate tumor volume according to the formula: $W^2 \times L/2$. The Student's t test and the Mann-Whitney U test, where applicable, were used to evaluate differences of tumor growth. All tests were two-sided with $\alpha=0.05$. The results show that Anti-158P1D7 mAbs are capable of retarding the growth of established human prostate carcinoma in mice (Figure 31).

These studies demonstrate a broad anti-tumor efficacy of anti-158P1D7 antibodies on initiation and progression of bladder cancer and prostate cancer and indicate that 158P1D7 antibodies to be efficacious in inhibiting and retarding the growth of 158P1D7-expressing tissues (Table I) in mouse models. Anti-158P1D7 antibodies inhibit tumor formation and retard the growth of already established tumors and prolong the survival of treated mice. Moreover, anti-158P1D7 mAbs demonstrate a dramatic inhibitory effect on the spread of local bladder tumor to distal sites, even in the presence of a large tumor burden. Thus, anti-158P1D7 mAbs are efficacious on major clinically relevant end points including lessened tumor growth, lessened metastasis, and prolongation of survival.

Example 37: Homology Comparison of 158P1D7 to Known Sequences

The 158P1D7 protein has 841 amino acids with calculated molecular weight of 95.1 kDa, and pI of 6.07. 158P1D7 is predicted to be a plasma membrane protein (0.46 PSORT http://psort.nibb.ac.jp/form.html) with a possibility of it being a nuclear protein (65% by PSORT http://psort.nibb.ac.jp/form2.html). 158P1D7 has a potential cleavage site between aa 626 and 627 and a potential signal site at aa 3-25.

158P1D7 contains a single transmembrane region from amino acids 611-633 with high probability that the amino-terminus resides outside, consistent with the topology of a Type 1 transmembrane protein (located on the World Wide Web at .cbs.dtu.dk/services/TMHMM). Also visualized is a short hydrophobic stretch from amino acids 3-25, consistent with the existence of an amino-terminal signal peptide. Based on the TMpred algorithm of Hofmann and Stoffel which utilizes TMBASE (K. Hofmann, W. Stoffel, TMBASE - A database of membrane spanning protein segments Biol. Chem. Hoppe-Seyler 374:166, 1993), 158P1D7 contains a primary transmembrane region from amino acids 609-633 and a secondary transmembrane region from amino acids 3-25 (contiguous amino acids with values greater than 0 on the plot have high probability of being transmembrane regions) with an orientation in which the amino terminus resides inside and the carboxyl terminus outside. An alternative model is also predicted that 158P1D7 is a Type 1 transmembrane protein in which the amino-terminus resides outside and the protein contains a secondary transmembrane domain signal peptide from amino acids 3-25 and a primary transmembrane domain from aa615-633. The transmembrane prediction algorithms are accessed through the ExPasy molecular biology server located on the World Wide Web at (.expasy.ch/tools/).

By use of the PubMed website of the N.C.B.I. located on the World Wide Web at (.ncbi.nlm.nih.gov/entrez) , it was found at the protein level that 158P1D7 shows best homology to the hypothetical protein FLJ22774 (PubMed record: gi 14149932) of unknown function, with 97% identity and 97% homology (Figure 4 and Figure 5A). The 158P1D7 protein demonstrates homology to a human protein similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit) (PubMed record: gi 6691962) with 36% identity and 52% homology (Figure 5B), to Slit proteins with 25% identity and 39% homology and to the leucine-rich repeat transmembrane family of proteins FLRT (Fibronectin-like domain-containing leucine-rich transmembrane protein), including FLRT2 with 26% identity and 43% homology, and FLRT3 with 34% identity and 53% homology.

Insulin-like growth factors (IGF) have been shown to play an important role in tumor growth including prostate, breast, brain and ovarian cancer (O'Brian et al, Urology. 2001, 58:1; Wang J et al Oncogene. 2001, 20:3857; Helle S et al, Br J Cancer. 2001, 85:74). IGFs produce their oncogenic effect by binding to specific cell surface receptors and activating survival as well as mitogenic pathways (Babajko S et al, Med Pediatr Oncol. 2001, 36:154; Scalia P et al, J Cell Biochem. 2001, 82:610). The activity of insulin-like growth factors is regulated by IGF binding proteins (IGF-BP) and the acid labile subunit (ALS) of IGF-BP (Zeslawski W et al, EMBO J. 2001, 20:3638; Jones JI. and Clemmons DR. Endocr. Rev. 1995, 16: 3). In the plasma, most IGFs exist as a ternary complex containing IGF-BP and ALS (Jones JI. and Clemmons DR. Endocr. Rev. 1995, 16: 3). Association with ALS allows the retention of the ternary complex in the vasculature and extends its

lifespan (Ueki I et al, Proc Natl Acad Sci U S A 2000, 97:6868). Studies in mice demonstrate the contribution of ALS to cell growth by showing that mice carrying mutant ALS exhibit a growth deficit (Ueki I et al, Proc Natl Acad Sci U S A 2000, 97:6868), indicating that ALS plays a critical role in the growth of tumor cells. The 158P1D7 protein serves as an IGF-ALS-like protein in that it facilitates the formation of the IGF ternary complex. The 158P1D7-induced IGF complex formation leads to increased growth of tumor cells expressing 158P1D7 which facilitates the growth of this malignancy *in vivo*. The induction of the IGF complex allows one to assay for monoclonal antibodies with neutralizing ability to disrupt, or enhancing capacity to help form, the ternary interaction.

Slit proteins were first identified in Drosophila as secreted proteins that regulate axon guidance and orientation (Rajagopalan S et al, Cell. 2000, 103:1033; Chen J et al, J Neurosci. 2001, 21:1548). Mammalian homologs were cloned in mice and humans, where they are shown to regulate migration and chemotaxis (Wu J et al, Nature. 2001, 410:948; Brose K and Tessier M, Curr Opin Neurobiol. 2001, 10:95). Slit proteins localize at two distinct subcellular sites within epithelial cells depending on cell stage, with Slit 3 predominantly localizing in the mitochondria and targeting to the cell surface in more confluent cells (Little MH et al, Am J Physiol Cell Physiol. 2001, 281:C486). The differential Slit localization suggests that Slit may function differently whether it is secreted, associated with the cell surface or retained in the mitochondria. The 158P1D7 protein functions as a Slit-like protein in that it binds to Roundabout receptors (Robos) on the surface of cells. 158P1D7 has homology (83% identity along entire length) with the murine Slitrk6 gene, a member of a new family of Leucine Rich Receptors (LRRs). The Slit family of LRRs is involved in neurite outgrowth and axonal guidance during development. These proteins also play a role in organ development by providing cues for branching morphogenesis in lung, kidney and other organs. The crystal structure for several LRRs has been determined. These proteins are shaped like a horseshoe with LRRs on both sides of a central flexible region. This horseshoe shape likely forms a central pocket where other proteins (binding partners) can interact. The term binding partner includes ligands, receptors, substrates, antibodies, and other molecules that interact with the 158P1D7 polypeptide through contact or proximity between particular portions of the binding partner and the 158P1D7 polypeptide. Binding partners for 158P1D7 polypeptides are expressed on both epithelial and mesenchymal cells within an organ. Known binding partners for the Slit family of LRRs include both the Robo family of genes and glypicans. Both of these potential protein interacting partners are aberrantly expressed in human cancers. Robos are Ig-like proteins that act as adhesion molecules. Interaction of specific Robo and Slit proteins results in cell migration with the ultimate outcome being either repulsion or attraction depending on intracellular signaling cascades. Mutations that disrupt interaction of Slit with Robo result in failure to repel migrating neurons during development. Moreover, mutations that disrupt functional interactions lead to organ failure and hyperproliferation in the developing lung. Mutational analysis has further shown that the LRR region is required for biologic activity of these receptors. 158P1D7 is overexpressed in a variety of human cancers including those derived from bladder and lung. Aberrant expression of this protein leads to enhanced cell growth, survival, increased metastasis and angiogenesis by disrupting or promoting protein interactions between 158P1D7 and specific binding partners on the surface of adjacent cells. Binding of 158P1D7 to Robo receptors (Robo-1, -2, -3 and -4) is observed in vitro, both as recombinant proteins and as cell surface molecules. Biological effects are induced when the Robo-1, -2, -3 or -4 receptors or glypican-binding partners binds to 158P1D7 on the cell surface. These activities are detected by adhesion, enhanced migration or repulsion in cell based assays. The interaction between 158P1D7 and Robo receptors leads to increased adhesion between 158P1D7-expressing tumor cells and endothelium or other cell types expressing Robo receptors, leading to spreading and metastasis of tumor cells as well as enhanced angiogenesis. Further, the association between 158P1D7 and Robo receptors allows one to screen for monoclonal antibodies with the ability to block (or enhance) the interaction in an *in vitro* assay. Such antibodies have a modulating effect on growth of 158P1D7 expressing tumors.

The FLRT (Fibronectin-like domain-containing leucine-rich transmembrane protein) family of transmembrane proteins has three members, FLRT1, FLRT2 and FLRT3, which contain 10 leucine-rich repeats flanked by cysteine-rich

domains, a fibronectin/collagen-like motif and an intracellular tail (Lacy SE et al, Genomics 1999, 62:417). Based on overall structure of the three proteins, a role in cell adhesion and receptor signaling is predicted. A Xenopus laevis ortholog of FLRT3 (XFLRT3) was identified that shows co-expression with FGFs (fibroblast growth factors) and is induced after activation and reduced following inhibition of signal transduction through the FGFs (Bottcher RT et al, Nature Cell Biol 2004, 6:38). The interaction between FGFRs (FGF receptors) and XFLRT3 indicates that XFLRT3 modulates FGF-induced signal transduction through the MAP kinase pathway. The 158P1D7 protein forms a complex with FGFRs that induces modulation of FGF-induced signal transduction through the MAP kinase (ERK-1 and ERK-2) pathway. FGF-induced signals are potentiated by expression of 158P1D7, which leads to an increase in the proliferative capacity of the cells. This significantly promotes unregulated growth of cancer cells expressing 158P1D7, contributing to their growth advantage *in vivo*. The interaction between 158P1D7 protein and FGFR allows one to screen for monoclonal antibodies with the ability to disrupt (or enhance) the association of these two molecules. Such antibodies have a modulating effect on growth of 158P1D7 expressing tumors.

Example 38: Identification and Confirmation of Signal Transduction Pathways

Many mammalian proteins have been reported to interact with signaling molecules and to participate in regulating signaling pathways. (J Neurochem. 2001; 76:217-223). In particular, IGF and IGF-BP have been shown to regulate mitogenic and survival pathways (Babajko S et al, Med Pediatr Oncol. 2001, 36:154; Scalia P et al, J Cell Biochem. 2001, 82:610). Using immunoprecipitation and Western blotting techniques, proteins are identified that associate with 158P1D7 and mediate signaling events. Several pathways known to play a role in cancer biology are regulated by 158P1D7, including phospholipid pathways such as PI3K, AKT, etc, adhesion and migration pathways, including FAK, Rho, Rac-1, etc, as well as mitogenic/survival cascades such as ERK, p38, etc. (Cell Growth Differ. 2000,11:279; J Biol Chem. 1999, 274:801; Oncogene. 2000, 19:3003, J. Cell Biol. 1997, 138:913.). Bioinformatic analysis revealed that 158P1D7 can become phosphorylated by serine/threonine as well as tyrosine kinases. Thus, the phosphorylation of 158P1D7 is provided by the present invention to lead to activation of the above listed pathways.

Using, e.g., Western blotting techniques, the ability of 158P1D7 to regulate these pathways is confirmed. Cells expressing or lacking 158P1D7 are either left untreated or stimulated with cytokines, hormones and anti-integrin antibodies. Cell lysates are analyzed using anti-phospho-specific antibodies (Cell Signaling, Santa Cruz Biotechnology) in order to detect phosphorylation and regulation of ERK, p38, AKT, PI3K, PLC and other signaling molecules. When 158P1D7 plays a role in the regulation of signaling pathways, whether individually or communally, it is used as a target for diagnostic, prognostic, preventative and therapeutic purposes.

To confirm that 158P1D7 directly or indirectly activates known signal transduction pathways in cells, luciferase (luc) based transcriptional reporter assays are carried out in cells expressing individual genes. These transcriptional reporters contain consensus-binding sites for known transcription factors that lie downstream of well-characterized signal transduction pathways. The reporters and examples of these associated transcription factors, signal transduction pathways, and activation stimuli are listed below:

1. NFkB-luc, NFkB/Rel; Ik-kinase/SAPK; growth/apoptosis/stress
2. SRE-luc, SRF/TCF/ELK1; MAPK/SAPK; growth/differentiation
3. AP-1-luc, FOS/JUN; MAPK/SAPK/PKC; growth/apoptosis/stress
4. ARE-luc, androgen receptor; steroids/MAPK; growth/differentiation/apoptosis
5. p53-luc, p53; SAPK; growth/differentiation/apoptosis
6. CRE-luc, CREB/ATF2; PKA/p38; growth/apoptosis/stress

Gene-mediated effects are assayed in cells showing mRNA expression. Luciferase reporter plasmids are introduced by lipid-mediated transfection (TFX-50, Promega). Luciferase activity, an indicator of relative transcriptional activity, is measured by incubation of cell extracts with luciferin substrate and luminescence of the reaction is monitored in a luminometer.

Signaling pathways activated by 158P1D7 are mapped and used for the identification and validation of therapeutic targets. When 158P1D7 is involved in cell signaling, it is used as target for diagnostic, prognostic, preventative and therapeutic purposes.

Example 39: Involvement in Tumor Progression

The 158P1D7 gene can contribute to the growth of cancer cells. The role of 158P1D7 in tumor growth is confirmed in a variety of primary and transfected cell lines including prostate, colon, bladder and kidney cell lines as well as NIH 3T3 cells engineered to stably express 158P1D7. Parental cells lacking 158P1D7 and cells expressing 158P1D7 are evaluated for cell growth using a well-documented proliferation assay (see, e.g., Fraser SP, Grimes JA, Djamgoz MB. Prostate. 2000;44:61, Johnson DE, Ochieng J, Evans SL. Anticancer Drugs. 1996, 7:288).

To confirm the role of 158P1D7 in the transformation process, its effect in colony forming assays is investigated. Parental NIH3T3 cells lacking 158P1D7 are compared to NHI-3T3 cells expressing 158P1D7, using a soft agar assay under stringent and more permissive conditions (Song Z. et al. Cancer Res. 2000, 60:6730).

To confirm the role of 158P1D7 in invasion and metastasis of cancer cells, a well-established assay is used, e.g., a Transwell Insert System assay (Becton Dickinson) (Cancer Res. 1999, 59:6010). Control cells, including prostate, colon, bladder and kidney cell lines lacking 158P1D7 are compared to cells expressing 158P1D7, respectively. Cells are loaded with the fluorescent dye, calcein, and plated in the top well of the Transwell insert coated with a basement membrane analog. Invasion is determined by fluorescence of cells in the lower chamber relative to the fluorescence of the entire cell population.

158P1D7 can also play a role in cell cycle and apoptosis. Parental cells and cells expressing 158P1D7 are compared for differences in cell cycle regulation using a well-established BrdU assay (Abdel-Malek ZA. J Cell Physiol. 1988, 136:247). In short, cells are grown under both optimal (full serum) and limiting (low serum) conditions are labeled with BrdU and stained with anti-BrdU Ab and propidium iodide. Cells are analyzed for entry into the G1, S, and G2M phases of the cell cycle. Alternatively, the effect of stress on apoptosis is evaluated in control parental cells and cells expressing 158P1D7, including normal and tumor bladder cells. Engineered and parental cells are treated with various chemotherapeutic agents, such as paclitaxel, gemcitabine, etc, and protein synthesis inhibitors, such as cycloheximide. Cells are stained with annexin V-FITC and cell death is measured by FACS analysis. The modulation of cell death by 158P1D7 can play a critical role in regulating tumor progression and tumor load.

When 158P1D7 plays a role in cell growth, transformation, invasion or apoptosis, it is used as a target for diagnostic, prognostic, preventative and therapeutic purposes.

Example 40: Involvement in Angiogenesis

Angiogenesis or new capillary blood vessel formation is necessary for tumor growth (Hanahan D, Folkman J. Cell. 1996, 86:353; Folkman J. Endocrinology. 1998 139:441). Several assays have been developed to measure angiogenesis *in vitro* and *in vivo*, such as the tissue culture assays, endothelial cell tube formation, and endothelial cell proliferation. Using these assays as well as *in vitro* neo-vascularization, the effect of 158P1D7 on angiogenesis is confirmed. For example, endothelial cells engineered to express 158P1D7 are evaluated using tube formation and proliferation assays. The effect of

158P1D7 is also confirmed in animal models *in vivo*. For example, cells either expressing or lacking 158P1D7 are implanted subcutaneously in immunocompromised mice. Endothelial cell migration and angiogenesis are evaluated 5-15 days later using immunohistochemistry techniques. When 158P1D7 affects angiogenesis, it is used as a target for diagnostic, prognostic, preventative and therapeutic purposes

Example 41: Regulation of Transcription

The above-indicated localization of 158P1D7 to the nucleus and its similarity to IGF-BP which has been found to activate signaling pathways and to regulate essential cellular functions, support the present invention use of 158P1D7 based on its role in the transcriptional regulation of eukaryotic genes. Regulation of gene expression is confirmed, e.g., by studying gene expression in cells expressing or lacking 158P1D7. For this purpose, two types of experiments are performed.

In the first set of experiments, RNA from parental and 158P1D7-expressing cells are extracted and hybridized to commercially available gene arrays (Clontech) (Smid-Koopman E et al. Br J Cancer. 2000. 83:246). Resting cells as well as cells treated with FBS or androgen are compared. Differentially expressed genes are identified in accordance with procedures known in the art. The differentially expressed genes are then mapped to biological pathways (Chen K et al., Thyroid. 2001. 11:41.).

In the second set of experiments, specific transcriptional pathway activation is evaluated using commercially available (e.g., Stratagene) luciferase reporter constructs including: NFkB-luc, SRE-luc, ELK1-luc, ARE-luc, p53-luc, and CRE-luc. These transcriptional reporters contain consensus binding sites for known transcription factors that lie downstream of well-characterized signal transduction pathways, and represent a good tool to ascertain pathway activation and screen for positive and negative modulators of pathway activation.

When 158P1D7 plays a role in gene regulation, it is used as a target for diagnostic, prognostic, preventative and therapeutic purposes.

Example 42: Subcellular Localization of 158P1D7

The cellular location of 158P1D7 is assessed using subcellular fractionation techniques widely used in cellular biology (Storrie B, et al. Methods Enzymol. 1990;182:203-25). A variety of cell lines, including prostate, kidney and bladder cell lines as well as cell lines engineered to express 158P1D7 are separated into nuclear, cytosolic and membrane fractions. Gene expression and location in nuclei, heavy membranes (lysosomes, peroxisomes, and mitochondria), light membranes (plasma membrane and endoplasmic reticulum), and soluble protein fractions are tested using Western blotting techniques.

Alternatively, 293T cells are transfected with an expression vector encoding individual genes, HIS-tagged (PCDNA 3.1 MYC/HIS, Invitrogen) and the subcellular localization of these genes is determined as described above. In short, the transfected cells are harvested and subjected to a differential subcellular fractionation protocol (Pemberton, P.A. et al, 1997, J of Histochemistry and Cytochemistry, 45:1697-1706). Location of the HIS-tagged genes is followed by Western blotting.

Using 158P1D7 antibodies, it is possible to demonstrate cellular localization by immunofluorescence and immunohistochemistry. For example, cells expressing or lacking 158P1D7 are adhered to a microscope slide and stained with anti-158P1D7 specific Ab. Cells are incubated with an FITC-coupled secondary anti-species Ab, and analyzed by fluorescent microscopy. Alternatively, cells and tissues lacking or expressing 158P1D7 are analyzed by IHC as described herein.

When 158P1D7 is localized to specific cell compartments, it is used as a target for diagnostic, preventative and therapeutic purposes.

Example 43: Involvement of 158P1D7 in Protein Trafficking.

Due to its similarity to Slit proteins, 158P1D7 can regulate intracellular trafficking and retention into mitochondrial and/or nuclear compartments. Its role in the trafficking of proteins can be confirmed using well-established methods (Valetti C. et al. Mol Biol Cell. 1999, 10:4107). For example, FITC-conjugated α2-macroglobulin is incubated with 158P1D7-expressing and 158P1D7-negative cells. The location and uptake of FITC-α2-macroglobulin is visualized using a fluorescent microscope. In another approach, the co-localization of 158P1D7 with vesicular proteins is confirmed by co-precipitation and Western blotting techniques and fluorescent microscopy.

Alternatively, 158P1D7-expressing and 158P1D7-lacking cells are compared using bodipy-ceramide labeled bovine serum albumine (Huber L et al. Mol. Cell. Biol. 1995, 15:918). Briefly, cells are allowed to take up the labeled BSA and are placed intermittently at 4°C and 18°C to allow for trafficking to take place. Cells are examined under fluorescent microscopy, at different time points, for the presence of labeled BSA in specific vesicular compartments, including Golgi, endoplasmic reticulum, etc.

In another embodiment, the effect of 158P1D7 on membrane transport is examined using biotin-avidin complexes. Cells either expressing or lacking 158P1D7 are transiently incubated with biotin. The cells are placed at 4°C or transiently warmed to 37°C for various periods of time. The cells are fractionated and examined by avidin affinity precipitation for the presence of biotin in specific cellular compartments. Using such assay systems, proteins, antibodies and small molecules are identified that modify the effect of 158P1D7 on vesicular transport. When 158P1D7 plays a role in intracellular trafficking, 158P1D7 is a target for diagnostic, prognostic, preventative and therapeutic purposes

Example 44: Protein-Protein Association

IGF and IGF-BP proteins have been shown to interact with other proteins, thereby forming protein complexes that can regulate protein localization, biological activity, gene transcription, and cell transformation (Zeslawski W et al, EMBO J. 2001, 20:3638; Yu H, Rohan T, J Natl Cancer Inst. 2000, 92:1472). Using immunoprecipitation techniques as well as two yeast hybrid systems, proteins are identified that associate with 158P1D7. Immunoprecipitates from cells expressing 158P1D7 and cells lacking 158P1D7 are compared for specific protein-protein associations.

Studies are performed to determine the extent of the association of 158P1D7 with receptors, such as the EGF and IGF receptors, and with intracellular proteins, such as IGF-BP, cytoskeletal proteins etc. Studies comparing 158P1D7 positive and 158P1D7 negative cells, as well as studies comparing unstimulated/resting cells and cells treated with epithelial cell activators, such as cytokines, growth factors and anti-integrin Ab reveal unique protein-protein interactions.

In addition, protein-protein interactions are confirmed using two yeast hybrid methodology (Curr Opin Chem Biol. 1999, 3:64). A vector carrying a library of proteins fused to the activation domain of a transcription factor is introduced into yeast expressing a 158P1D7-DNA-binding domain fusion protein and a reporter construct. Protein-protein interaction is detected by colorimetric reporter activity. Specific association with surface receptors and effector molecules directs one of skill to the mode of action of 158P1D7, and thus identifies therapeutic, prognostic, preventative and/or diagnostic targets for cancer. This and similar assays are also used to identify and screen for small molecules that interact with 158P1D7.

When 158P1D7 associates with proteins or small molecules it is used as a target for diagnostic, prognostic, preventative and therapeutic purposes.

Example 45: Transcript Variants of 158P1D7

Transcript variants are variants of mature mRNA from the same gene which arise by alternative transcription or alternative splicing. Alternative transcripts are transcripts from the same gene but start transcription at different points. Splice

variants are mRNA variants spliced differently from the same transcript. In eukaryotes, when a multi-exon gene is transcribed from genomic DNA, the initial RNA is spliced to produce functional mRNA, which has only exons and is used for translation into an amino acid sequence. Accordingly, a given gene can have zero to many alternative transcripts and each transcript can have zero to many splice variants. Each transcript variant has a unique exon makeup, and can have different coding and/or non-coding (5' or 3' end) portions, from the original transcript. Transcript variants can code for similar or different proteins with the same or a similar function or can encode proteins with different functions, and can be expressed in the same tissue at the same time, or in different tissues at the same time, or in the same tissue at different times, or in different tissues at different times. Proteins encoded by transcript variants can have similar or different cellular or extracellular localizations, e.g., secreted versus intracellular.

Transcript variants are identified by a variety of art-accepted methods. For example, alternative transcripts and splice variants are identified by full-length cloning experiment, or by use of full-length transcript and EST sequences. First, all human ESTs were grouped into clusters which show direct or indirect identity with each other. Second, ESTs in the same cluster were further grouped into sub-clusters and assembled into a consensus sequence. The original gene sequence is compared to the consensus sequence(s) or other full-length sequences. Each consensus sequence is a potential splice variant for that gene (see, e.g., URL www.doubletwist.com/products/c11_agentsOverview.jhtml). Even when a variant is identified that is not a full-length clone, that portion of the variant is very useful for antigen generation and for further cloning of the full-length splice variant, using techniques known in the art.

Moreover, computer programs are available in the art that identify transcript variants based on genomic sequences. Genomic-based transcript variant identification programs include FgenesH (A. Salamov and V. Solovyev, "Ab initio gene finding in Drosophila genomic DNA," Genome Research. 2000 April;10(4):516-22); Grail (URL compbio.ornl.gov/Grail-bin/EmptyGrailForm) and GenScan (URL genes.mit.edu/GENSCAN.html). For a general discussion of splice variant identification protocols see., e.g., Southan, C., A genomic perspective on human proteases, FEBS Lett. 2001 Jun 8; 498(2-3):214-8; de Souza, S.J., *et al.*, Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags, Proc. Natl Acad Sci U S A. 2000 Nov 7; 97(23):12690-3.

To further confirm the parameters of a transcript variant, a variety of techniques are available in the art, such as full-length cloning, proteomic validation, PCR-based validation, and 5' RACE validation, etc. (see e.g., Proteomic Validation: Brennan, S.O., *et al.*, Albumin banks peninsula: a new termination variant characterized by electrospray mass spectrometry, Biochem Biophys Acta. 1999 Aug 17;1433(1-2):321-6; Ferranti P, *et al.*, Differential splicing of pre-messenger RNA produces multiple forms of mature caprine alpha(s1)-casein, Eur J Biochem. 1997 Oct 1;249(1):1-7. For PCR-based Validation: Wellmann S, *et al.*, Specific reverse transcription-PCR quantification of vascular endothelial growth factor (VEGF) splice variants by LightCycler technology, Clin Chem. 2001 Apr;47(4):654-60; Jia, H.P., *et al.*, Discovery of new human beta-defensins using a genomics-based approach, Gene. 2001 Jan 24; 263(1-2):211-8. For PCR-based and 5' RACE Validation: Brigle, K.E., *et al.*, Organization of the murine reduced folate carrier gene and identification of variant splice forms, Biochem Biophys Acta. 1997 Aug 7; 1353(2): 191-8).

It is known in the art that genomic regions are modulated in cancers. When the genomic region to which a gene maps is modulated in a particular cancer, the alternative transcripts or splice variants of the gene are modulated as well. Disclosed herein is that 158P1D7 has a particular expression profile related to cancer. Alternative transcripts and splice variants of 158P1D7 may also be involved in cancers in the same or different tissues, thus serving as tumor-associated markers/antigens.

Using the full-length gene and EST sequences, four transcript variants were identified, designated as 158P1D7 v.3, v.4, v.5 and v.6. The boundaries of the exon in the original transcript, 158P1D7 v.1 were shown in Table BILL-I. Compared with 158P1D7 v.1, transcript variant 158P1D7 v.3 has spliced out 2069-2395 from variant 158P1D7 v.1, as shown in Figure

12. Variant 158P1D7 v.4 spliced out 1162-2096 of variant 158P1D7 v.1. Variant 158P1D7 v.5 added one exon to the 5' and extended 2 bp to the 5' end and 288 bp to the 3' end of variant 158P1D7 v.1. Theoretically, each different combination of exons in spatial order, e.g. exon 1 of v.5 and exons 1 and 2 of v.3 or v.4, is a potential splice variant.

The variants of 158P1D7 include those that lack a transmembrane motif, but include a signal peptide indicating that they are secreted proteins (v.4 and v.6). Secreted proteins such as v.4 and v.6 serve as biomarkers of cancer existence and progression. The levels of such variant proteins in the serum of cancer patients serves as a prognostic marker of cancer disease or its progression, particularly of cancers such as those listed in Table I. Moreover, such secreted proteins are targets of monoclonal antibodies and related binding molecules. Accordingly, secreted proteins such as these serve as targets for diagnostics, prognostics, prophylactics and therapeutics for human malignancies. Targeting of secreted variants of 158P1D7 is particularly preferred when they have pathogy-related or cancer-related effects on cells/tissues.

Tables LI (a)-(d) through LIV(a)-(d) are set forth on a variant-by-variant bases. Tables LI(a)-(d) shows nucleotide sequence of the transcript variant. Tables LII(a)-(d) shows the alignment of the transcript variant with nucleic acid sequence of 158P1D7 v.1. Tables LIII (a)-(d) lays out amino acid translation of the transcript variant for the identified reading frame orientation. Tables LIV(a)-(d) displays alignments of the amino acid sequence encoded by the splice variant with that of 158P1D7 v.1.

Example 46: Single Nucleotide Polymorphisms of 158P1D7

A Single Nucleotide Polymorphism (SNP) is a single base pair variation in a nucleotide sequence at a specific location. At any given point of the genome, there are four possible nucleotide base pairs: A/T, C/G, G/C and T/A. Genotype refers to the specific base pair sequence of one or more locations in the genome of an individual. Haplotype refers to the base pair sequence of more than one location on the same DNA molecule (or the same chromosome in higher organisms), often in the context of one gene or in the context of several tightly linked genes. SNP that occurs on a cDNA is called cSNP. This cSNP may change amino acids of the protein encoded by the gene and thus change the functions of the protein. Some SNP cause inherited diseases; others contribute to quantitative variations in phenotype and reactions to environmental factors including diet and drugs among individuals. Therefore, SNP and/or combinations of alleles (called haplotypes) have many applications, including diagnosis of inherited diseases, determination of drug reactions and dosage, identification of genes responsible for diseases, and analysis of the genetic relationship between individuals (P. Nowotny, J. M. Kwon and A. M. Goate, " SNP analysis to dissect human traits," Curr. Opin. Neurobiol. 2001 Oct; 11(5):637-641; M. Pirmohamed and B. K. Park, "Genetic susceptibility to adverse drug reactions," Trends Pharmacol. Sci. 2001 Jun; 22(6):298-305; J. H. Riley, C. J. Allan, E. Lai and A. Roses, " The use of single nucleotide polymorphisms in the isolation of common disease genes," Pharmacogenomics. 2000 Feb; 1(1):39-47; R. Judson, J. C. Stephens and A. Windemuth, "The predictive power of haplotypes in clinical response," Pharmacogenomics. 2000 feb; 1(1):15-26).

SNP are identified by a variety of art-accepted methods (P. Bean, "The promising voyage of SNP target discovery," Am. Clin. Lab. 2001 Oct-Nov; 20(9):18-20; K. M. Weiss, "In search of human variation," Genome Res. 1998 Jul; 8(7):691-697; M. M. She, "Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies," Clin. Chem. 2001 Feb; 47(2):164-172). For example, SNP can be identified by sequencing DNA fragments that show polymorphism by gel-based methods such as restriction fragment length polymorphism (RFLP) and denaturing gradient gel electrophoresis (DGGE). They can also be discovered by direct sequencing of DNA samples pooled from different individuals or by comparing sequences from different DNA samples. With the rapid accumulation of sequence data in public and private databases, one can discover SNP by comparing sequences using computer programs (Z. Gu, L. Hillier and P. Y. Kwok, "Single nucleotide polymorphism hunting in cyberspace," Hum. Mutat. 1998; 12(4):221-225). SNP can be verified and genotype or haplotype of an individual can be determined by a variety of methods including direct sequencing

and high throughput microarrays (P. Y. Kwok, "Methods for genotyping single nucleotide polymorphisms," Annu. Rev. Genomics Hum. Genet. 2001; 2:235-258; M. Kokoris, K. Dix, K. Moynihan, J. Mathis, B. Erwin, P. Grass, B. Hines and A. Duesterhoeft, "High-throughput SNP genotyping with the Masscode system," Mol. Diagn. 2000 Dec; 5(4):329-340).

Using the methods described above, one SNP was identified in the original transcript, 158P1D7 v.1, at positions 1546 (A/G). The transcripts or proteins with alternative allele was designated as variant 158P1D7 v.2. Figure 17 shows the schematic alignment of the SNP variants. Figure 18 shows the schematic alignment of protein variants, corresponding to nucleotide variants. Nucleotide variants that code for the same amino acid sequence as v.1 are not shown in Figure 18. These alleles of the SNP, though shown separately here, can occur in different combinations (haplotypes) and in any one of the transcript variants (such as 158P1D7 v.5) that contains the site of the SNP.

Example 47: Therapeutic and Diagnostic use of Anti-158P1D7 Antibodies in Humans.

Anti-158P1D7 monoclonal antibodies are safely and effectively used for diagnostic, prophylactic, prognostic and/or therapeutic purposes in humans. Western blot and immunohistochemical analysis of cancer tissues and cancer xenografts with anti-158P1D7 mAb show strong extensive staining in carcinoma but significantly lower or undetectable levels in normal tissues. Detection of 158P1D7 in carcinoma and in metastatic disease demonstrates the usefulness of the mAb as a diagnostic and/or prognostic indicator. Anti-158P1D7 antibodies are therefore used in diagnostic applications such as immunohistochemistry of kidney biopsy specimens to detect cancer from suspect patients.

As determined by flow cytometry, anti-158P1D7 mAb specifically binds to carcinoma cells. Thus, anti-158P1D7 antibodies are used in diagnostic whole body imaging applications, such as radioimmunoscintigraphy and radioimmunotherapy, (see, e.g., Potamianos S., et. al. Anticancer Res 20(2A):925-948 (2000)) for the detection of localized and metastatic cancers that exhibit expression of 158P1D7. Shedding or release of an extracellular domain of 158P1D7 into the extracellular milieu, such as that seen for alkaline phosphodiesterase B10 (Meerson, N. R., Hepatology 27:563-568 (1998)), allows diagnostic detection of 158P1D7 by anti-158P1D7 antibodies in serum and/or urine samples from suspect patients.

Anti-158P1D7 antibodies that specifically bind 158P1D7 are used in therapeutic applications for the treatment of cancers that express 158P1D7. Anti-158P1D7 antibodies are used as an unconjugated modality and as conjugated form in which the antibodies are attached to one of various therapeutic or imaging modalities well known in the art, such as a prodrugs, enzymes or radioisotopes. In preclinical studies, unconjugated and conjugated anti-158P1D7 antibodies are tested for efficacy of tumor prevention and growth inhibition in the SCID mouse cancer xenograft models, e.g., kidney cancer models AGS-K3 and AGS-K6, (see, e.g., the Example entitled "158P1D7 Monoclonal Antibody-mediated Inhibition of Bladder and Lung Tumors *In Vivo*"). Either conjugated and unconjugated anti-158P1D7 antibodies are used as a therapeutic modality in human clinical trials either alone or in combination with other treatments as described in following Examples.

Example 48: Human Clinical Trials for the Treatment and Diagnosis of Human Carcinomas through use of Human Anti-158P1D7 Antibodies *In vivo*

Antibodies are used in accordance with the present invention which recognize an epitope on 158P1D7, and are used in the treatment of certain tumors such as those listed in Table I. Based upon a number of factors, including 158P1D7 expression levels, tumors such as those listed in Table I are presently preferred indications. In connection with each of these indications, three clinical approaches are successfully pursued.

I.) Adjunctive therapy: In adjunctive therapy, patients are treated with anti-158P1D7 antibodies in combination with a chemotherapeutic or antineoplastic agent and/or radiation therapy. Primary cancer targets, such as those

listed in Table I, are treated under standard protocols by the addition anti-158P1D7 antibodies to standard first and second line therapy. Protocol designs address effectiveness as assessed by reduction in tumor mass as well as the ability to reduce usual doses of standard chemotherapy. These dosage reductions allow additional and/or prolonged therapy by reducing dose-related toxicity of the chemotherapeutic agent. Anti-158P1D7 antibodies are utilized in several adjunctive clinical trials in combination with the chemotherapeutic or antineoplastic agents adriamycin (advanced prostrate carcinoma), cisplatin (advanced head and neck and lung carcinomas), taxol (breast cancer), and doxorubicin (preclinical).

II.) Monotherapy: In connection with the use of the anti-158P1D7 antibodies in monotherapy of tumors, the antibodies are administered to patients without a chemotherapeutic or antineoplastic agent. In one embodiment, monotherapy is conducted clinically in end stage cancer patients with extensive metastatic disease. Patients show some disease stabilization. Trials demonstrate an effect in refractory patients with cancerous tumors.

III.) Imaging Agent: Through binding a radionuclide (e.g., iodine or yttrium (I^{131}, Y^{90}) to anti-158P1D7 antibodies, the radiolabeled antibodies are utilized as a diagnostic and/or imaging agent. In such a role, the labeled antibodies localize to both solid tumors, as well as, metastatic lesions of cells expressing 158P1D7. In connection with the use of the anti-158P1D7 antibodies as imaging agents, the antibodies are used as an adjunct to surgical treatment of solid tumors, as both a pre-surgical screen as well as a post-operative follow-up to determine what tumor remains and/or returns. In one embodiment, a (^{111}In)-158P1D7 antibody is used as an imaging agent in a Phase I human clinical trial in patients having a carcinoma that expresses 158P1D7 (by analogy see, *e.g.*, Divgi *et al. J. Natl. Cancer Inst.* 83:97-104 (1991)). Patients are followed with standard anterior and posterior gamma camera. The results indicate that primary lesions and metastatic lesions are identified.

Dose and Route of Administration

As appreciated by those of ordinary skill in the art, dosing considerations can be determined through comparison with the analogous products that are in the clinic. Thus, anti-158P1D7 antibodies can be administered with doses in the range of 5 to 400 mg/m^2, with the lower doses used, e.g., in connection with safety studies. The affinity of anti-158P1D7 antibodies relative to the affinity of a known antibody for its target is one parameter used by those of skill in the art for determining analogous dose regimens. Further, anti-158P1D7 antibodies that are fully human antibodies, as compared to the chimeric antibody, have slower clearance; accordingly, dosing in patients with such fully human anti-158P1D7 antibodies can be lower, perhaps in the range of 50 to 300 mg/m^2, and still remain efficacious. Dosing in mg/m^2, as opposed to the conventional measurement of dose in mg/kg, is a measurement based on surface area and is a convenient dosing measurement that is designed to include patients of all sizes from infants to adults.

Three distinct delivery approaches are useful for delivery of anti-158P1D7 antibodies. Conventional intravenous delivery is one standard delivery technique for many tumors. However, in connection with tumors in the peritoneal cavity, such as tumors of the ovaries, biliary duct, other ducts, and the like, intraperitoneal administration may prove favorable for obtaining high dose of antibody at the tumor and to also minimize antibody clearance. In a similar manner, certain solid tumors possess vasculature that is appropriate for regional perfusion. Regional perfusion allows for a high dose of antibody at the site of a tumor and minimizes short term clearance of the antibody.

Clinical Development Plan (CDP)

Overview: The CDP follows and develops treatments of anti-158P1D7 antibodies in connection with adjunctive therapy, monotherapy, and as an imaging agent. Trials initially demonstrate safety and thereafter confirm efficacy in repeat doses. Trails are open label comparing standard chemotherapy with standard therapy plus anti-158P1D7 antibodies. As will be appreciated, one criteria that can be utilized in connection with enrollment of patients is 158P1D7 expression levels in their tumors as determined by biopsy.

As with any protein or antibody infusion-based therapeutic, safety concerns are related primarily to (i) cytokine release syndrome, i.e., hypotension, fever, shaking, chills; (ii) the development of an immunogenic response to the material (i.e., development of human antibodies by the patient to the antibody therapeutic, or HAHA response); and, (iii) toxicity to normal cells that express 158P1D7. Standard tests and follow-up are utilized to monitor each of these safety concerns. Anti-158P1D7 antibodies are found to be safe upon human administration.

Example 49: Human Clinical Trial Adjunctive Therapy with Human Anti-158P1D7 Antibody and Chemotherapeutic Agent

A phase I human clinical trial is initiated to assess the safety of six intravenous doses of a human anti-158P1D7 antibody in connection with the treatment of a solid tumor, e.g., a cancer of a tissue listed in Table I. In the study, the safety of single doses of anti-158P1D7 antibodies when utilized as an adjunctive therapy to an antineoplastic or chemotherapeutic agent as defined herein, such as, without limitation: cisplatin, topotecan, doxorubicin, adriamycin, taxol, or the like, is assessed. The trial design includes delivery of six single doses of an anti-158P1D7 antibody with dosage of antibody escalating from approximately about 25 mg/m ^{2} to about 275 mg/m ^{2} over the course of the treatment in accordance with the following schedule:

	Day 0	Day 7	Day 14	Day 21	Day 28	Day 35
mAb Dose	25 mg/m ^{2}	75 mg/m ^{2}	125 mg/m ^{2}	175 mg/m ^{2}	225 mg/m ^{2}	275 mg/m ^{2}
Chemotherapy (standard dose)	+	+	+	+	+	+

Patients are closely followed for one-week following each administration of antibody and chemotherapy. In particular, patients are assessed for the safety concerns mentioned above: (i) cytokine release syndrome, i.e., hypotension, fever, shaking, chills; (ii) the development of an immunogenic response to the material (i.e., development of human antibodies by the patient to the human antibody therapeutic, or HAHA response); and, (iii) toxicity to normal cells that express 158P1D7. Standard tests and follow-up are utilized to monitor each of these safety concerns. Patients are also assessed for clinical outcome, and particularly reduction in tumor mass as evidenced by MRI or other imaging.

The anti-158P1D7 antibodies are demonstrated to be safe and efficacious, Phase II trials confirm the efficacy and refine optimum dosing.

Example 50: Human Clinical Trial: Monotherapy with Human Anti-158P1D7 Antibody

Anti-158P1D7 antibodies are safe in connection with the above-discussed adjunctive trial, a Phase II human clinical trial confirms the efficacy and optimum dosing for monotherapy. Such trial is accomplished, and entails the same safety and outcome analyses, to the above-described adjunctive trial with the exception being that patients do not receive chemotherapy concurrently with the receipt of doses of anti-158P1D7 antibodies.

Example 51: Human Clinical Trial: Diagnostic Imaging with Anti-158P1D7 Antibody

Once again, as the adjunctive therapy discussed above is safe within the safety criteria discussed above, a human clinical trial is conducted concerning the use of anti-158P1D7 antibodies as a diagnostic imaging agent. The protocol is

designed in a substantially similar manner to those described in the art, such as in Divgi *et al. J. Natl. Cancer Inst.* 83:97-104 (1991). The antibodies are found to be both safe and efficacious when used as a diagnostic modality.

Example 52: RNA Interfernece (RNAi)

RNA interference (RNAi) technology is implemented to a variety of cell assays relevant to oncology. RNAi is a post-transcriptional gene silencing mechanism activated by double-stranded RNA (dsRNA). RNAi induces specific mRNA degradation leading to changes in protein expression and subsequently in gene function. In mammalian cells, these dsRNAs called short interfering RNA (siRNA) have the correct composition to activate the RNAi pathway targeting for degradation, specifically some mRNAs. *See,* Elbashir S.M., *et. al.*, Duplexes of 21-nucleotide RNAs Mediate RNA interference in Cultured Mammalian Cells, Nature 411(6836):494-8 (2001). Thus, RNAi technology is used successfully in mammalian cells to silence targeted genes.

Loss of cell proliferation control is a hallmark of cancerous cells; thus, assessing the role of 158P1D7 in cell survival/proliferation assays is relevant. Accordingly, RNAi was used to investigate the function of the 158P1D7 antigen. To generate siRNA for 158P1D7, algorithms were used that predict oligonucleotides that exhibit the critical molecular parameters (G:C content, melting temperature, etc.) and have the ability to significantly reduce the expression levels of the 158P1D7 protein when introduced into cells. Accordingly, one targeted sequence for the 158P1D7 siRNA is: 5' AAGCTCATTCTAGCGGGAAAT 3' (SEQ ID NO: 42)(oligo 158P1D7.b). In accordance with this Example, 158P1D7 siRNA compositions are used that comprise siRNA (double stranded, short interfering RNA) that correspond to the nucleic acid ORF sequence of the 158P1D7 protein or subsequences thereof. Thus, siRNA subsequences are used in this manner are generally 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30,31, 32, 33, 34, 35 or more than 35 contiguous RNA nucleotides in length. These siRNA sequences are complementary and non-complementary to at least a portion of the mRNA coding sequence. In a preferred embodiment, the subsequences are 19-25 nucleotides in length, most preferably 21-23 nucleotides in length. In preferred embodiments, these siRNA achieve knockdown of 158P1D7 antigen in cells expressing the protein and have functional effects as described below.

The selected siRNA (158P1D7.b oligo) was tested in numerous cell lines in the survival/proliferation MTS assay (measures cellular metabolic activity). Tetrazolium-based colorimetric assays (i.e., MTS) detect viable cells exclusively, since living cells are metabolically active and therefore can reduce tetrazolium salts to colored formazan compounds; dead cells, however do not. Moreover, this 158P1D7.b oligo achieved knockdown of 158P1D7 antigen in cells expressing the protein and had functional effects as described below using the following protocols.

Mammalian siRNA transfections: The day before siRNA transfection, the different cell lines were plated in media (RPMI 1640 with 10% FBS w/o antibiotics) at $2x10^3$ cells/well in 80 µl (96 well plate format) for the survival/MTS assay. In parallel with the 158P1D7 specific siRNA oligo, the following sequences were included in every experiment as controls: a) Mock transfected cells with Lipofectamine 2000 (Invitrogen, Carlsbad, CA) and annealing buffer (no siRNA); b) Luciferase-4 specific siRNA (targeted sequence: 5'-AAGGGACGAAGACGAACACUUCTT-3') (SEQ ID NO: 43); and, c) Eg5 specific siRNA (targeted sequence: 5'-AACTGAAGACCTGAAGACAATAA-3') (SEQ ID NO: 44). SiRNAs were used at 10nM and 1µg/ml Lipofectamine 2000 final concentration.

The procedure was as follows: The siRNAs were first diluted in OPTIMEM (serum-free transfection media, Invitrogen) at 0.1uM µM (10-fold concentrated) and incubated 5-10 min RT. Lipofectamine 2000 was diluted at 10 µg/ml (10-fold concentrated) for the total number transfections and incubated 5-10 minutes at room temperature (RT). Appropriate amounts of diluted 10-fold concentrated Lipofectamine 2000 were mixed 1:1 with diluted 10-fold concentrated siRNA and

incubated at RT for 20-30" (5-fold concentrated transfection solution). 20 µls of the 5-fold concentrated transfection solutions were added to the respective samples and incubated at 37°C for 96 hours before analysis.

MTS assay: The MTS assay is a colorimetric method for determining the number of viable cells in proliferation, cytotoxicity or chemosensitivity assays based on a tetrazolium compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS(b)] and an electron coupling reagent (phenazine ethosulfate; PES). Assays were performed by adding a small amount of the Solution Reagent directly to culture wells, incubating for 1-4 hours and then recording absorbance at 490nm with a 96-well plate reader. The quantity of colored formazan product as measured by the amount of 490nm absorbance is directly proportional to the mitochondrial activity and/or the number of living cells in culture.

In order to address the function of 158P1D7 in cells, 158P1D7 was silenced by transfecting the endogenously expressing 158P1D7 cell lines (LNCaP and PC3) with the 158P1D7 specific siRNA (158P1D7.b) along with negative siRNA controls (Luc4, targeted sequence not represented in the human genome) and a positive siRNA control (targeting Eg5) (Figure 29). The results indicated that when these cells are treated with siRNA specifically targeting the 158P1D7 mRNA, the resulting "158P1D7 deficient cells" showed diminished cell viability or proliferation as measured by this assay (see oligo 158P1D7.b treated cells). This effect is likely caused by an active induction of apoptosis. The reduced viability is measured by the increased release (and activity) of a mitochondrial enzyme that occurs predominantly in apoptotic cells.

As control, 3T3 cells, a cell line with no detectable expression of 158P1D7 mRNA, was also treated with the panel of siRNAs (including oligo 158P1D7.b) and no phenotype was observed. This result reflects the fact that the specific protein knockdown in the LNCaP and PC3 cells is not a function of general toxicity, since the 3T3 cells did not respond to the 158P1D7.b oligo. The differential response of the three cell lines to the Eg5 control is a reflection of differences in levels of cell transfection and responsiveness of the cell lines to oligo treatment (Figure 29).

Together, these data indicate that 158P1D7 plays an important role in the proliferation of cancer cells and that the lack of 158P1D7 clearly decreases the survival potential of these cells. It is to be noted that 158P1D7 is constitutively expressed in many tumor cell lines. 158P1D7 serves a role in malignancy; it expression is a primary indicator of disease, where such disease is often characterized by high rates of uncontrolled cell proliferation and diminished apoptosis. Correlating cellular phenotype with gene knockdown following RNAi treatments is important, and allows one to draw valid conclusions and rule out toxicity or other non-specific effects of these reagents. To this end, assays to measure the levels of expression of both protein and mRNA for the target after RNAi treatments are important, including Western blotting, FACS staining with antibody, immunoprecipitation, Northern blotting or RT-PCR (Taqman or standard methods). Any phenotypic effect of the siRNAs in these assays should be correlated with the protein and/or mRNA knockdown levels in the same cell lines. Knockdown of 158P1D7 is achieved using the 158P1D7.b oligo as measured by Western blotting and RT-PCR analysis.

A method to analyze 158P1D7 related cell proliferation is the measurement of DNA synthesis as a marker for proliferation. Labeled DNA precursors (i.e. ^{3}H-Thymidine) are used and their incorporation to DNA is quantified. Incorporation of the labeled precursor into DNA is directly proportional to the amount of cell division occurring in the culture. Another method used to measure cell proliferation is performing clonogenic assays. In these assays, a defined number of cells are plated onto the appropriate matrix and the number of colonies formed after a period of growth following siRNA treatment is counted.

In 158P1D7 cancer target validation, complementing the cell survival/proliferation analysis with apoptosis and cell cycle profiling studies are considered. The biochemical hallmark of the apoptotic process is genomic DNA fragmentation, an irreversible event that commits the cell to die. A method to observe fragmented DNA in cells is the immunological detection of histone-complexed DNA fragments by an immunoassay (i.e. cell death detection ELISA) which measures the enrichment

of histone-complexed DNA fragments (mono- and oligo-nucleosomes) in the cytoplasm of apoptotic cells. This assay does not require pre-labeling of the cells and can detect DNA degradation in cells that do not proliferate in vitro (i.e. freshly isolated tumor cells).

The most important effector molecules for triggering apoptotic cell death are caspases. Caspases are proteases that when activated cleave numerous substrates at the carboxy-terminal site of an aspartate residue mediating very early stages of apoptosis upon activation. All caspases are synthesized as pro-enzymes and activation involves cleavage at aspartate residues. In particular, caspase 3 seems to play a central role in the initiation of cellular events of apoptosis. Assays for determination of caspase 3 activation detect early events of apoptosis. Following RNAi treatments, Western blot detection of active caspase 3 presence or proteolytic cleavage of products (i.e. PARP) found in apoptotic cells further support an active induction of apoptosis. Because the cellular mechanisms that result in apoptosis are complex, each has its advantages and limitations. Consideration of other criteria/endpoints such as cellular morphology, chromatin condensation, membrane blebbing, apoptotic bodies help to further support cell death as apoptotic. Since not all the gene targets that regulate cell growth are anti-apoptotic, the DNA content of permeabilized cells is measured to obtain the profile of DNA content or cell cycle profile. Nuclei of apoptotic cells contain less DNA due to the leaking out to the cytoplasm (sub-G1 population). In addition, the use of DNA stains (i.e., propidium iodide) also differentiate between the different phases of the cell cycle in the cell population due to the presence of different quantities of DNA in G0/G1, S and G2/M. In these studies the subpopulations can be quantified.

For the 158P1D7 gene, RNAi studies facilitate the understanding of the contribution of the gene product in cancer pathways. Such active RNAi molecules have use in identifying assays to screen for mAbs that are active anti-tumor therapeutics. Further, siRNA are administered as therapeutics to cancer patients for reducing the malignant growth of several cancer types, including those listed in Table 1. When 158P1D7 plays a role in cell survival, cell proliferation, tumorigenesis, or apoptosis, it is used as a target for diagnostic, prognostic, preventative and/or therapeutic purposes

Example 53: 158P1D7 Functional Assays

I. Enhanced proliferation and cell cycle modulation in 158P1D7 expressing cells.

Enhanced proliferation and entry into S-phase of tumor cells relative to normal cells is a hallmark of the cancer cell phenotype. To address the effect of expression of 158P1D7 on the proliferation rate of normal cells, two rodent cell lines (3T3 and Rat-1) were infected with virus containing the 158P1D7 gene and stable cells expressing 158P1D7 antigen were derived, as well as empty vector control cells expressing the selection marker neomycin (Neo). The cells were grown overnight in 0.5% FBS and then compared to cells treated with 10% FBS. The cells were evaluated for proliferation at 18-96 hr post-treatment by a ^{3}H-thymidine incorporation assay and for cell cycle analysis by a BrdU incorporation/propidium iodide staining assay. The results in Figure 32 show that the Rat-1 cells expressing the 158P1D7 antigen grew effectively in low serum concentrations (0.1%) compared to the Rat-1-Neo cells. Similar results were obtained for the 3T3 cells expressing 158P1D7 versus Neo only. To assess cell proliferation by another methodology, the cells were stained with BrdU and propidium iodide. Briefly, cells were labeled with 10 ▯M BrdU, washed, trypsinized and fixed in 0.4% paraformaldehyde and 70% ethanol. Anti-BrdU-FITC (Pharmigen) was added to the cells, the cells were washed and then incubated with 10 ▯g/ml propidium iodide for 20 min prior to washing and analysis for fluorescence at 488 nm. The results in Figure 33 show that there was increased labeling of cells in S-phase (DNA synthesis phase of the cell cycle) in 3T3 cells that expressed the 158P1D7 antigen relative to control cells. These results confirm those measured by ^{3}H-thymidine incorporation, and indicate that cells that express 158P1D7 antigen have an enhanced proliferative capacity and survive in low serum conditions. Accordingly, 158P1D7 expressing cells have increased potential for growth as tumor cells in vivo.

II. Recombinant extracellular domain (ECD) binding to cell surface.

Cell-cell interactions are essential in maintaining tissue/organ integrity and homeostasis, both of which become deregulated during tumor formation and progression. Additionally, cell-cell interactions facilitate tumor cell attachment during metastasis and activation of endothelium for increased angiogenesis. To address interaction between the gene product of 158P1D7 and a putative ligand, an assay was established to identify the interaction between the extracellular domain (ECD) (amino acids 16-608) of 158P1D7 antigen and primary endothelium. Human umbilical vein endothelial cells (HUVEC) were grown in 0.1% FBS in media for 3 hr. Cells were washed, detached in 10 mM EDTA and resuspended in 10% FBS. Recombinant 158P1D7 ECD (described in Example entitled "Production of Recombinant 158P1D7 in Eukaryotic Systems") was added to cells, and the cells were washed prior to the addition of MAb M15/X68.2.22 at 1 ug/ml. After washing, secondary Ab (anti-mouse-PE, 1:400) was added to cells for 1 hr on ice. Cells were washed and fixed in 1% formalin for 3 hr on ice, then resuspended in PBS and analyzed by flow cytometry. Figure 26A shows that the158P1D7 ECD bound directly to the surface of HUVEC cells as detected by the 158P1D7 specific MAb. In a similar embodiment, recombinant ECD of 158P1D7 was iodinated to high specific activity using the iodogen (1,3,4,5-tetrachloro-3a,6a-diphenylglycoluril) method. HUVEC cells at 90% confluency in 6 well plates were incubated with 1 nM of ^{125}I-158P1D7 ECD in the presence (non-specific binding) or absence (Total binding) of 50 fold excess unlabeled ECD for 2 hours at either 4°C or 37°C. Cells were washed, solubilized in 0.5M NaOH, and subjected to gamma counting. The data in Figure 26B shows specific binding of 158P1D7 ECD to HUVEC cells suggesting the presence of a 158P1D7 receptor on HUVEC cells. These results indicate that 158P1D7 antigen is involved in cell-cell interactions that facilitate tumor growth, activation of endothelium for tumor vascularization or tumor cell metastasis. The data also indicate that 158P1D7 antigen shed from the cell surface of expressing cells may bind to cells in an autocrine or paracrine fashion to induce cell effector functions.

Example 54: Detection of 158P1D7 protein in cancer patient specimens using Immunohistochemistry.

To determine the expression of 158P1D7 protein, specimens were obtained from various cancer patients and stained using an affinity purified monoclonal antibody raised against the peptide encoding amino acids 274-285 of 158P1D7 (See the Example Entitled "Generation of 158P1D7 Monoclonal Antibodies (mAbs)"), formalin fixed, paraffin embedded tissues were cut into 4 micron sections and mounted on glass slides. The sections were dewaxed, rehydrated and treated with antigen retrieval solution (Antigen Retrieval Citra Solution; BioGenex, 4600 Norris Canyon Road, San Ramon, CA, 94583) at high temperature. Sections were then incubated in mouse monoclonal anti-158P1D7 antibody, M15-68(2)22, for 3 hours. The slides were washed three times in buffer and further incubated with DAKO EnVision+™ peroxidase-conjugated goat anti-mouse immunoglobulin secondary antibody (DAKO Corporation, Carpenteria, CA) for 1 hour. The sections were then washed in buffer, developed using the DAB kit (SIGMA Chemicals), counterstained using hematoxylin, and analyzed by bright field microscopy. The results showed expression of 158P1D7 in cancer patients' tissue (Figure 36). Generally, in bladder transitional cell carcinoma expression of 158P1D7 was mainly around the cell membrane indicating that 158P1D7 is membrane associated in these tissues. 49.3% of bladder transitional cell carcinoma samples tested were positive for 158P1D7 (Table LVIII).

These results indicate that 158P1D7 is a target for diagnostic, prophylactic, prognostic and therapeutic applications in cancer.

The present invention is not to be limited in scope by the embodiments disclosed herein, which are intended as single illustrations of individual aspects of the invention, and any that are functionally equivalent are within the scope of the invention. Various modifications to the models and methods of the invention, in addition to those described herein, will become apparent to those skilled in the art from the foregoing description and teachings, and are similarly intended to fall

within the scope of the invention. Such modifications or other embodiments can be practiced without departing from the true scope and spirit of the invention.

All documents and publications recited herein are hereby incorporated in their entirety as if fully set forth.

TABLES:

TABLE I: Tissues that Express 158P1D7 When Malignant

Bladder, Prostate, Colon, Lung, Breast, Ovary, Skin, Cervix

TABLE II: AMINO ACID ABBREVIATIONS

SINGLE LETTER	THREE LETTER	FULL NAME
F	Phe	phenylalanine
L	Leu	leucine
S	Ser	serine
Y	Tyr	tyrosine
C	Cys	cysteine
W	Trp	tryptophan
P	Pro	proline
H	His	histidine
Q	Gln	glutamine
R	Arg	arginine
I	Ile	isoleucine
M	Met	methionine
T	Thr	threonine
N	Asn	asparagine
K	Lys	lysine
V	Val	valine
A	Ala	alanine
D	Asp	aspartic acid
E	Glu	glutamic acid
G	Gly	glycine

TABLE III: AMINO ACID SUBSTITUTION MATRIX

Adapted from the GCG Software 9.0 BLOSUM62 amino acid substitution matrix (block substitution matrix). The higher the value, the more likely a substitution is found in related, natural proteins. (See world wide web URL ikp.unibe.ch/manual/blosum62.html)

A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	.
4	0	-2	-1	-2	0	-2	-1	-1	-1	-1	-2	-1	-1	-1	1	0	0	-3	-2	A
	9	-3	-4	-2	-3	-3	-1	-3	-1	-1	-3	-3	-3	-3	-1	-1	-1	-2	-2	C
		6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-3	D
			5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-2	E
				6	-3	-1	0	-3	0	0	-3	-4	-3	-3	-2	-2	-1	1	3	F
					6	-2	-4	-2	-4	-3	0	-2	-2	-2	0	-2	-3	-2	-3	G
						8	-3	-1	-3	-2	1	-2	0	0	-1	-2	-3	-2	2	H
							4	-3	2	1	-3	-3	-3	-3	-2	-1	3	-3	-1	I
								5	-2	-1	0	-1	1	2	0	-1	-2	-3	-2	K
									4	2	-3	-3	-2	-2	-2	-1	1	-2	-1	L
										5	-2	-2	0	-1	-1	-1	1	-1	-1	M
											6	-2	0	0	1	0	-3	-4	-2	N
												7	-1	-2	-1	-1	-2	-4	-3	P
													5	1	0	-1	-2	-2	-1	Q
														5	-1	-1	-3	-3	-2	R
															4	1	-2	-3	-2	S
																5	0	-2	-2	T
																	4	-3	-1	V
																		11	2	W
																			7	Y

TABLE IV

HLA Class I/II Motifs/Supermotifs

TABLE IV (A): HLA Class I Supermotifs/Motifs

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	TI*LVMS*		**FWY**
A2	LIVM*ATQ*		IV*MATL*
A3	VSMA*TLI*		**RK**
A24	YF*WIVLMT*		FI*YWLM*
B7	**P**		VILF*MWYA*
B27	**RHK**		FYL*WMIVA*
B44	E*D*		**FWYLIMVA**
B58	**ATS**		**FWY***LIVMA*
B62	**QL***IVMP*		**FWYMIVLA**
MOTIFS			
A1	**TSM**		**Y**
A1		DE*AS*	**Y**
A2.1	LM*VQIAT*		V*LIMAT*
A3	LMVISATF*CGD*		KYR*HFA*
A11	VTMLISAGN*CDF*		K*RYH*
A24	YFW*M*		**FLIW**
A*3101	MVT*ALIS*		R*K*
A*3301	MVALF*IST*		**RK**
A*6801	AVT*MSLI*		**RK**
B*0702	**P**		LMF*WYAIV*
B*3501	**P**		LMFWY*IVA*
B51	**P**		LIVF*WYAM*
B*5301	**P**		IMFWY*ALV*
B*5401	**P**		ATIV*LMFWY*

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE IV (B): HLA CLASS II SUPERMOTIF

1	6	9
W, F, Y, V, .I, L	A, V, I, L, P, C, S, T	A, V, I, L, C, S, T, M, Y

TABLE IV (C) HLA Class II Motifs

MOTIFS		1° anchor 1	2	3	4	5	1° anchor 6	7	8	9
DR4	preferred	FMY*LIVW*	M	T		I	VST*CPALIM*	MH		MH
	deleterious				W			R		WDE
DR1	preferred	MF*LIVWY*			PAMQ		VMAT*SPLIC*	M		AVM
	deleterious		C	CH	FD	CWD		GDE	D	
DR7	preferred	MF*LIVWY*	M	W	A		IVMSA*CTPL*	M		IV
	deleterious		C		G			GRD	N	G
DR3	MOTIFS	1° anchor 1	2	3	1° anchor 4	5	1° anchor 6			
Motif a preferred		LIVMFY			D					
Motif b preferred		LIVMFAY			DNQEST		KRH			
DR Supermotif		MF*LIVWY*					VMSTA*CPLI*			

Italicized residues indicate less preferred or "tolerated" residues

TABLE IV (D) HLA Class I Supermotifs

SUPER-MOTIFS	POSITION:	1	2	3	4	5	6	7	8	C-terminus
A1			1° Anchor TI*LVMS*							1° Anchor FWY
A2			1° Anchor LIVM*ATQ*							1° Anchor LIVMAT
A3	Preferred		1° Anchor VSMA*TLI*	YFW (4/5)			YFW (3/5)	YFW (4/5)	P (4/5)	1° Anchor RK
	deleterious	DE (3/5); P (5/5)		DE (4/5)						
A24			1° Anchor YF*WIVLMT*							1° Anchor FIY*WLM*
B7	Preferred	FWY (5/5) LIVM (3/5)	1° Anchor P	FWY (4/5)					FWY (3/5)	1°Anchor VILF*MWYA*
	deleterious	DE (3/5); P(5/5); G(4/5); A(3/5); QN(3/5)				DE (3/5)	G (4/5)	QN (4/5)	DE (4/5)	
B27			1° Anchor RHK							1°Anchor FYL*WMIVA*
B44			1° Anchor E*D*							1° Anchor FWYLIMVA
B58			1° Anchor ATS							1° Anchor FWY*LIVMA*
B62			1° Anchor QL*IVMP*							1° Anchor FWY*MIVLA*

Italicized residues indicate less preferred or "tolerated" residues

TABLE IV (E) HLA Class I Motifs

	POSITION 1		2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A1 9-mer	preferred	GFYW	1°Anchor STM	DEA	YFW		P	DEQN	YFW	1°Anchor Y	
	deleterious	DE		RHKLIVMP	A	G	A				
A1 9-mer	preferred	GRHK	ASTCLIVM	1°Anchor DEAS	GSTC		ASTC	LIVM	DE	1°Anchor Y	
	deleterious	A	RHKDEPYFW		DE	PQN	RHK	PG	GP		
A1 10-mer	preferred	YFW	1°Anchor STM	DEAQN	A	YFWQN		PASTC	GDE	P	1°Anchor Y
	deleterious	GP		RHKGLIVM	DE	RHK	QNA	RHKYFW	RHK	A	
A1 10-mer	preferred	YFW	STCLIVM	1°Anchor DEAS	A	YFW		PG	G	YFW	1°Anchor Y
	deleterious	RHK	RHKDEPYFW			P	G		PRHK	QN	
A2.1 9-mer	preferred	YFW	1°Anchor LM*IVQAT*	YFW	STC	YFW		A	P	1°Anchor *VLIMAT*	
	deleterious	DEP		DERKH			RKH	DERKH			
	POSITION: 1		2	3	4	5	6	7	8	9	C-Terminus
A2.1 10-mer	preferred	AYFW	1°Anchor LM*IVQAT*	LVIM	G		*G*		FYWL VIM		1°Anchor *VLIMAT*
	deleterious	DEP		DE	RKHA	P		RKH	DERKH	RKH	
A3	preferred	RHK	1°Anchor LMVISATFCGD	YFW	PRHKYFW	A	YFW		P	1°Anchor KYR*HFA*	
	deleterious	DEP		DE							
A11	preferred	A	1°Anchor VTLMISAGN*CD F*	YFW	YFW	A	YFW	YFW	P	1°Anchor *KRYH*	
	deleterious	DEP						A	G		
A24 9-mer	preferred	YFWRHK	1°Anchor YFW*M*		STC			YFW	YFW	1°Anchor FLIW	
	deleterious	DEG		DE	G	QNP	DERHKG		AQN		
A24 10-mer	Preferred		1°Anchor YFW*M*		P	YFWP		P			1°Anchor FLIW
	Deleterious			GDE	QN	RHK	DE	A	QN	DEA	
A3101	Preferred	RHK	1°Anchor MVT*ALIS*	YFW	P		YFW	YFW	AP	1°Anchor R*K*	
	Deleterious	DEP		DE		ADE	DE	DE	DE		
A3301	Preferred		1°Anchor MVALF*IST*	YFW				AYFW		1°Anchor RK	
	Deleterious	GP		DE							
A6801	Preferred	YFWSTC	1°Anchor AVT*MSLI*			YFWLIVM		YFW	P	1°Anchor RK	
	deleterious	GP		DEG		RHK			A		
B0702	Preferred	RHKFWY	1°Anchor P	RHK		RHK	RHK	RHK	PA	1°Anchor LMF*WYAI V*	
	deleterious	DEQNP		DEP	DE	DE	GDE	QN	DE		
B3501	Preferred	FWYLIVM	1°Anchor P	FWY				FWY		1°Anchor LMFWY*IV*	

	POSITION	1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A1 9-mer	preferred	GFYW	1°Anchor STM	DEA	YFW		P	DEQN	YFW	1°Anchor Y	
	deleterious	DE		RHKLIVMP	A	G	A				
A1 9-mer	preferred	GRHK	ASTCLIVM	1°Anchor DE*AS*	GSTC		ASTC	LIVM	DE	1°Anchor Y	
	deleterious	A	RHKDEPYFW		DE	PQN	RHK	PG	GP		
	deleterious	AGP				G	G			*A*	
B51	Preferred	LIVMFWY	1°Anchor P	FWY	STC	FWY		G	FWY	1°Anchor LIVF*WYA M*	
	deleterious	AGPDER HKSTC				DE	G	DEQN	GDE		
B5301	preferred	LIVMFWY	1°Anchor P	FWY	STC	FWY		LIVMFWY	FWY	1°Anchor IMFWY*AL V*	
	deleterious	AGPQN					G	RHKQN	DE		
B5401	preferred	FWY	1°Anchor P	FWYLIVM		LIVM		ALIVM	FWYA P	1°Anchor ATIV*LMF WY*	
	deleterious	GPQNDE		GDESTC		RHKDE	DE	QNDGE	DE		

(Italicized residues indicate less preferred or "tolerated" residues. The information in this Table is specific for 9-mers unless otherwise specified.)

TABLE IV (F):

Summary of HLA-supertypes								
Overall phenotypic frequencies of HLA-supertypes in different ethnic populations								
Specificity			Phenotypic frequency					
Supertype	Position 2	C-Terminus	Caucasian	N.A. Black	Japanese	Chinese	Hispanic	Average
B7	P	AILMVFWY	43.2	55.1	57.1	43.0	49.3	49.5
A3	AILMVST	RK	37.5	42.1	45.8	52.7	43.1	44.2
A2	AILMVT	AILMVT	45.8	39.0	42.4	45.9	43.0	42.2
A24	YF (WIVLMT)	FI (YWLM)	23.9	38.9	58.6	40.1	38.3	40.0
B44	E (D)	FWYLIMVA	43.0	21.2	42.9	39.1	39.0	37.0
A1	TI (LVMS)	FWY	47.1	16.1	21.8	14.7	26.3	25.2
B27	RHK	FYL (WMI)	28.4	26.1	13.3	13.9	35.3	23.4
B62	QL (IVMP)	FWY (MIV)	12.6	4.8	36.5	25.4	11.1	18.1
B58	ATS	FWY (LIV)	10.0	25.1	1.6	9.0	5.9	10.3

TABLE IV (G):

Calculated population coverage afforded by different HLA-supertype combinations						
HLA-supertypes	Phenotypic frequency					
	Caucasian	N.A Blacks	Japanese	Chinese	Hispanic	Average
A2, A3 and B7	83.0	86.1	87.5	88.4	86.3	86.2
A2, A3, B7, A24, B44 and A1	99.5	98.1	100.0	99.5	99.4	99.3
A2, A3, B7, A24, B44, A1, B27, B62, and B 58	99.9	99.6	100.0	99.8	99.9	99.8
Motifs indicate the residues defining supertype specificites. The motifs incorporate residues determined on the basis of published data to be recognized by multiple alleles within the supertype. Residues within brackets are additional residues also predicted to be tolerated by multiple alleles within the supertype.						

Tables V-XVIII:

Table V-V1-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
150	VIEPSAFSK	900.000
436	NLEYLYLEY	225.000
812	LVEQTKNEY	45.000
828	HAEPDYLEV	45.000
711	GSDAKHLQR	37.500
546	CTSPGHLDK	25.000
265	SICPTPPVY	10.000
351	NIESLSDLR	9.000
799	LMETLMYSR	9.000
173	ESLPPNIFR	7.500
650	DNSPVHLQY	6.250
601	LTDAVPLSV	6.250
174	SLPPNIFRF	5.000
100	IADIEIGAF	5.000
682	MVSPMVHVY	5.000
102	DIEIGAFNG	4.500
134	GLENLEFLQ	4.500
47	NCEAKGIKM	4.500
383	LVEYFTLEM	4.500
401	VLEEGSFMN	4.500
388	TLEMLHLGN	4.500
749	FQDASSLYR	3.750
56	VSEISVPPS	2.700
561	NSEILCPGL	2.700
431	FLGLHNLEY	2.500
291	INDSRMSTK	2.500
142	QADNNFITV	2.500
502	ILDDLDLLT	2.500
522	SCDLVGLQQ	2.500
223	NCDLLQLKT	2.500
771	ITEYLRKNI	2.250
232	WLENMPPQS	1.800
171	AIESLPPNI	1.800
137	NLEFLQADN	1.800
355	LSDLRPPPQ	1.500
380	KSDLVEYFT	1.500
59	ISVPPSRPF	1.500
255	GSILSRLKK	1.500
540	VTDDILCTS	1.250
308	TKAPGLIPY	1.250

Table V-V1-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
817	KNEYFELKA	1.125
743	STEFLSFQD	1.125
359	RPPPQNPRK	1.000
246	VCNSPPFFK	1.000
417	YLNGNHLTK	1.000
433	GLHNLEYLY	1.000
785	DMEAHYPGA	0.900
398	RIEVLEEGS	0.900
701	EEEEERNEK	0.900
833	YLEVLEQQT	0.900
513	DLEDNPWDC	0.900
123	SLEILKEDT	0.900
203	FLEHIGRIL	0.900
36	NCEEKDGTM	0.900
699	HLEEEEERN	0.900
214	QLEDNKWAC	0.900
573	PSMPTQTSY	0.750
81	TNDFSGLTN	0.625
192	GNQLQTLPY	0.625
301	TSILKLPTK	0.600
631	LVLHRRRRY	0.500
643	QVDEQMRDN	0.500
610	LILGLLIMF	0.500
407	FMNLTRLQK	0.500
89	NAISIHLGF	0.500
187	HLDLRGNQL	0.500
511	QIDLEDNPW	0.500
627	GIVVLVLHR	0.500
472	QVLPPHIFS	0.500
593	TADTILRSL	0.500
337	VLSPSGLLI	0.500
210	ILDLQLEDN	0.500
615	LIMFITIVF	0.500
473	VLPPHIFSG	0.500
730	LTGSNMKYK	0.500
447	IKEILPGTF	0.450
669	TTERPSASL	0.450
441	YLEYNAIKE	0.450
802	TLMYSRPRK	0.400
683	VSPMVHVYR	0.300
547	TSPGHLDKK	0.300

Table V-V1-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
32	DSLCNCEEK	0.300
723	EQENHSPLT	0.270
276	HEDPSGSLH	0.250
769	LGITEYLRK	0.250
76	LTMLHTNDF	0.250
235	NMPPQSIIG	0.250
196	QTLPYVGFL	0.250
738	KTTNQSTEF	0.250
372	GNIIHSLMK	0.250
287	ATSSINDSR	0.250
551	HLDKKELKA	0.250
825	ANLHAEPDY	0.250
148	ITVIEPSAF	0.250
729	PLTGSNMKY	0.250
584	VTTPATTTN	0.250
664	KTTHHTTER	0.250
526	VGLQQWIQK	0.250
801	ETLMYSRPR	0.250
297	STKTTSILK	0.250

Table V-V3-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
7	HMGAHEELK	0.100
2	SLYEQHMGA	0.050
3	LYEQHMGAH	0.045
1	ASLYEQHMG	0.015
8	MGAHEELKL	0.013
6	QHMGAHEEL	0.001
5	EQHMGAHEE	0.000
4	YEQHMGAHE	0.000

Table V-V4-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is

9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	Subsequence	Score
3	HSLMKSILW	0.075
8	SILWSKASG	0.020
11	WSKASGRGR	0.015
7	KSILWSKAS	0.015
9	ILWSKASGR	0.010
5	LMKSILWSK	0.010
1	IIHSLMKSI	0.010
4	SLMKSILWS	0.005
12	SKASGRGRR	0.005
13	KASGRGRRE	0.001
6	MKSILWSKA	0.001
2	IHSLMKSIL	0.001
14	ASGRGRREE	0.000
10	LWSKASGRG	0.000

Table VI-V1-HLA-A1-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	Subsequence	Score
56	VSEISVPPSR	27.000
669	TTERPSASLY	11.250
210	ILDLQLEDNK	10.000
781	QLQPDMEAHY	10.000
150	VIEPSAFSKL	9.000
171	AIESLPPNIF	9.000
828	HAEPDYLEVL	9.000
123	SLEILKEDTF	9.000
398	RIEVLEEGSF	9.000
812	LVEQTKNEYF	9.000
173	ESLPPNIFRF	7.500
546	CTSPGHLDKK	5.000
134	GLENLEFLQA	4.500
401	VLEEGSFMNL	4.500
380	KSDLVEYFTL	3.750
456	NPMPKLKVLY	2.500
505	DLDLLTQIDL	2.500
502	ILDDLDLLTQ	2.500
743	STEFLSFQDA	2.250
771	ITEYLRKNIA	2.250
682	MVSPMVHVYR	2.000
214	QLEDNKWACN	1.800
355	LSDLRPPPQN	1.500

Table VI-V1-HLA-A1-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	Subsequence	Score
264	ESICPTPPVY	1.500
753	SSLYRNILEK	1.500
561	NSEILCPGLV	1.350
601	LTDAVPLSVL	1.250
276	HEDPSGSLHL	1.250
590	TTNTADTILR	1.250
149	TVIEPSAFSK	1.000
106	GAFNGLGLLK	1.000
801	ETLMYSRPRK	1.000
545	LCTSPGHLDK	1.000
824	KANLHAEPDY	1.000
525	LVGLQQWIQK	1.000
300	TTSILKLPTK	1.000
477	HIFSGVPLTK	1.000
100	IADIEIGAFN	1.000
768	QLGITEYLRK	1.000
245	VVCNSPPFFK	1.000
721	LLEQENHSPL	0.900
700	LEEEEERNEK	0.900
102	DIEIGAFNGL	0.900
441	YLEYNAIKEI	0.900
436	NLEYLYLEYN	0.900
36	NCEEKDGTML	0.900
513	DLEDNPWDCS	0.900
383	LVEYFTLEML	0.900
388	TLEMLHLGNN	0.900
137	NLEFLQADNN	0.900
232	WLENMPPQSI	0.900
47	NCEAKGIKMV	0.900
747	LSFQDASSLY	0.750
711	GSDAKHLQRS	0.750
723	EQENHSPLTG	0.675
728	SPLTGSNMKY	0.625
830	EPDYLEVLEQ	0.625
435	HNLEYLYLEY	0.625
191	RGNQLQTLPY	0.625
643	QVDEQMRDNS	0.500
223	NCDLLQLKTW	0.500
142	QADNNFITVI	0.500
60	SVPPSRPFQL	0.500
765	ELQQLGITEY	0.500

Table VI-V1-HLA-A1-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	Subsequence	Score
609	VLILGLLIMF	0.500
453	GTFNPMPKLK	0.500
630	VLVLHRRRRY	0.500
42	GTMLINCEAK	0.500
472	QVLPPHIFSG	0.500
593	TADTILRSLT	0.500
337	VLSPSGLLIH	0.500
811	VLVEQTKNEY	0.500
187	HLDLRGNQLQ	0.500
614	LLIMFITIVF	0.500
603	DAVPLSVLIL	0.500
200	YVGFLEHIGR	0.500
522	SCDLVGLQQW	0.500
203	FLEHIGRILD	0.450
759	ILEKERELQQ	0.450
706	RNEKEGSDAK	0.450
785	DMEAHYPGAH	0.450
351	NIESLSDLRP	0.450
439	YLYLEYNAIK	0.400
59	ISVPPSRPFQ	0.300
727	HSPLTGSNMK	0.300
419	NGNHLTKLSK	0.250
310	APGLIPYITK	0.250
681	HMVSPMVHVY	0.250
783	QPDMEAHYPG	0.250
432	LGLHNLEYLY	0.250
119	INHNSLEILK	0.250
451	LPGTFNPMPK	0.250
371	AGNIIHSLMK	0.250
254	KGSILSRLKK	0.250
796	ELKLMETLMY	0.250
584	VTTPATTTNT	0.250
820	YFELKANLHA	0.225
817	KNEYFELKAN	0.225
793	AHEELKLMET	0.225
358	LRPPPQNPRK	0.200

Table VI-V3-HLA-A1-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	**Subsequence**	**Score**
2	ASLYEQHMGA	0.075
8	HMGAHEELKL	0.025
1	SASLYEQHMG	0.010
7	QHMGAHEELK	0.010
3	SLYEQHMGAH	0.010
4	LYEQHMGAHE	0.009
9	MGAHEELKLM	0.003
6	EQHMGAHEEL	0.002
5	YEQHMGAHEE	0.000

Table VI-V4-HLA-A1-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	**Subsequence**	**Score**
9	SILWSKASGR	0.100
4	HSLMKSILWS	0.075
8	KSILWSKASG	0.030
5	SLMKSILWSK	0.020
12	WSKASGRGRR	0.015
1	NIIHSLMKSI	0.010
2	IIHSLMKSIL	0.010
3	IHSLMKSILW	0.003
10	ILWSKASGRG	0.001
14	KASGRGRREE	0.001
11	LWSKASGRGR	0.001
6	LMKSILWSKA	0.001
7	MKSILWSKAS	0.001
13	SKASGRGRRE	0.000

Table VII-V1-HLA-A2-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
465	YLNNNLLQV	735.860
614	LLIMFITIV	423.695

Table VII-V1-HLA-A2-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
193	NQLQTLPYV	330.059
616	IMFITIVFC	285.492
140	FLQADNNFI	263.950
415	KLYLNGNHL	239.259
439	YLYLEYNAI	230.356
611	ILGLLIMFI	224.357
2	KLWIHLFYS	158.832
429	GMFLGLHNL	131.296
581	YLMVTTPAT	126.833
463	VLYLNNNLL	116.211
574	SMPTQTSYL	84.856
71	LLNNGLTML	83.527
4	WIHLFYSSL	77.017
305	KLPTKAPGL	74.768
613	GLLIMFITI	73.343
213	LQLEDNKWA	71.445
826	NLHAEPDYL	57.572
803	LMYSRPRKV	54.652
501	NILDDLDLL	50.218
798	KLMETLMYS	50.051
527	GLQQWIQKL	49.134
158	KLNRLKVLI	36.515
178	NIFRFVPLT	33.135
225	DLLQLKTWL	32.604
462	KVLYLNNNL	24.206
767	QQLGITEYL	21.597
116	QLHINHNSL	21.362
68	QLSLLNNGL	21.362
502	ILDDLDLLT	20.776
70	SLLNNGLTM	18.382
470	LLQVLPPHI	17.736
391	MLHLGNNRI	17.736
164	VLILNDNAI	17.736
337	VLSPSGLLI	17.736
774	YLRKNIAQL	17.177
450	ILPGTFNPM	16.047
323	QLPGPYCPI	15.649
367	KLILAGNII	14.971
316	YITKPSTQL	13.512
141	LQADNNFIT	12.523
214	QLEDNKWAC	9.777

Table VII-V1-HLA-A2-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
582	LMVTTPATT	9.149
758	NILEKEREL	8.912
17	SLHSQTPVL	8.759
182	FVPLTHLDL	8.598
609	VLILGLLIM	7.964
295	RMSTKTTSI	7.535
309	KAPGLIPYI	6.415
539	TVTDDILCT	6.149
618	FITIVFCAA	5.970
596	TILRSLTDA	5.813
432	LGLHNLEYL	5.437
479	FSGVPLTKV	4.804
517	NPWDCSCDL	4.745
544	ILCTSPGHL	4.721
531	WIQKLSKNT	4.713
597	ILRSLTDAV	4.403
524	DLVGLQQWI	4.304
290	SINDSRMST	4.201
681	HMVSPMVHV	3.928
425	KLSKGMFLG	3.479
608	SVLILGLLI	3.378
336	KVLSPSGLL	3.147
147	FITVIEPSA	3.142
48	CEAKGIKMV	3.111
722	LEQENHSPL	2.895
16	ISLHSQTPV	2.856
99	NIADIEIGA	2.801
163	KVLILNDNA	2.758
92	SIHLGFNNI	2.726
400	EVLEEGSFM	2.720
384	VEYFTLEML	2.547
442	LEYNAIKEI	2.538
302	SILKLPTKA	2.527
453	GTFNPMPKL	2.525
154	SAFSKLNRL	2.525
45	LINCEAKGI	2.439
393	HLGNNRIEV	2.365
624	CAAGIVVLV	2.222
833	YLEVLEQQT	2.194
455	FNPMPKLKV	2.088
621	IVFCAAGIV	2.040

Table VII-V1-HLA-A2-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
408	MNLTRLQKL	2.017
646	EQMRDNSPV	1.957
481	GVPLTKVNL	1.869
780	AQLQPDMEA	1.864
196	QTLPYVGFL	1.805
604	AVPLSVLIL	1.763
473	VLPPHIFSG	1.690
487	VNLKTNQFT	1.683
675	ASLYEQHMV	1.680
612	LGLLIMFIT	1.674
821	FELKANLHA	1.644
175	LPPNIFRFV	1.613
494	FTHLPVSNI	1.533
474	LPPHIFSGV	1.466
709	KEGSDAKHL	1.454
620	TIVFCAAGI	1.435

Table VII-V3-HLA-A2-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
2	SLYEQHMGA	65.180
8	MGAHEELKL	0.237
6	QHMGAHEEL	0.027
1	ASLYEQHMG	0.002
4	YEQHMGAHE	0.001
7	HMGAHEELK	0.000
5	EQHMGAHEE	0.000
3	LYEQHMGAH	0.000

Table VII-V4-HLA-A2-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
1	IIHSLMKSI	5.609
4	SLMKSILWS	3.488
9	ILWSKASGR	0.210
8	SILWSKASG	0.038
6	MKSILWSKA	0.020
5	LMKSILWSK	0.011
2	IHSLMKSIL	0.010
7	KSILWSKAS	0.002
13	KASGRGRRE	0.000
3	HSLMKSILW	0.000
14	ASGRGRREE	0.000
11	WSKASGRGR	0.000
12	SKASGRGRR	0.000
10	LWSKASGRG	0.000

Table VII-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
613	GLLIMFITIV	922.161
431	FLGLHNLEYL	609.108
616	IMFITIVFCA	301.064
600	SLTDAVPLSV	285.163
417	YLNGNHLTKL	226.014
473	VLPPHIFSGV	224.653
70	SLLNNGLTML	181.794
433	GLHNLEYLYL	176.240
166	ILNDNAIESL	167.806
407	FMNLTRLQKL	163.232
174	SLPPNIFRFV	145.364
425	KLSKGMFLGL	142.060
581	YLMVTTPATT	126.833
409	NLTRLQKLYL	117.493
610	LILGLLIMFI	114.142
746	FLSFQDASSL	98.267
213	LQLEDNKWAC	97.424
141	LQADNNFITV	93.387
465	YLNNNLLQVL	92.666
369	ILAGNIIHSL	83.527
415	KLYLNGNHLT	83.462
140	FLQADNNFIT	81.516
158	KLNRLKVLIL	70.507
611	ILGLLIMFIT	69.289
78	MLHTNDFSGL	69.001

Table VII-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
615	LIMFITIVFC	54.353
802	TLMYSRPRKV	51.468
531	WIQKLSKNTV	43.992
469	NLLQVLPPHI	38.601
67	FQLSLLNNGL	36.864
803	LMYSRPRKVL	34.412
115	KQLHINHNSL	28.049
462	KVLYLNNNLL	24.206
86	GLTNAISIHL	21.362
401	VLEEGSFMNL	18.106
44	MLINCEAKGI	17.736
596	TILRSLTDAV	17.338
621	IVFCAAGIVV	15.695
501	NILDDLDLLT	15.544
4	WIHLFYSSLL	13.512
486	KVNLKTNQFT	12.552
163	KVLILNDNAI	11.822
336	KVLSPSGLLI	11.822
60	SVPPSRPFQL	10.841
282	SLHLAATSSI	10.433
110	GLGLLKQLHI	10.433
766	LQQLGITEYL	9.923
126	ILKEDTFHGL	9.902
15	CISLHSQTPV	9.563
582	LMVTTPATTT	9.149
257	ILSRLKKESI	8.691
517	NPWDCSCDLV	7.571
568	GLVNNPSMPT	7.452
441	YLEYNAIKEI	7.064
295	RMSTKTTSIL	6.326
678	YEQHMVSPMV	6.221
195	LQTLPYVGFL	6.055
770	GITEYLRKNI	5.881
322	TQLPGPYCPI	5.871
382	DLVEYFTLEM	5.805
192	GNQLQTLPYV	5.743
374	IIHSLMKSDL	4.993
647	QMRDNSPVHL	4.807
623	FCAAGIVVLV	4.804
305	KLPTKAPGLI	4.747
263	KESICPTPPV	4.733

Table VII-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
457	PMPKLKVLYL	4.294
428	KGMFLGLHNL	4.153
2	KLWIHLFYSS	4.113
656	LQYSMYGHKT	4.110
574	SMPTQTSYLM	3.588
227	LQLKTWLENM	3.571
343	LLIHCQERNI	3.547
490	KTNQFTHLPV	3.381
220	WACNCDLLQL	3.139
232	WLENMPPQSI	3.071
738	KTTNQSTEFL	2.799
555	KELKALNSEI	2.627
721	LLEQENHSPL	2.324
390	EMLHLGNNRI	2.091
328	YCPIPCNCKV	2.088
212	DLQLEDNKWA	2.049
526	VGLQQWIQKL	2.017
605	VPLSVLILGL	2.017
798	KLMETLMYSR	1.820
313	LIPYITKPST	1.742
577	TQTSYLMVTT	1.738
380	KSDLVEYFTL	1.698
204	LEHIGRILDL	1.624
198	LPYVGFLEHI	1.587
608	SVLILGLLIM	1.517
108	FNGLGLLKQL	1.475
6	HLFYSSLLAC	1.437
488	NLKTNQFTHL	1.421
814	EQTKNEYFEL	1.413
825	ANLHAEPDYL	1.391
512	IDLEDNPWDC	1.335
818	NEYFELKANL	1.329
575	MPTQTSYLMV	1.158
77	TMLHTNDFSG	1.155

Table VIII-V3-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
8	HMGAHEELKL	0.525
3	SLYEQHMGAH	0.292
9	MGAHEELKLM	0.127
2	ASLYEQHMGA	0.120
6	EQHMGAHEEL	0.080
5	YEQHMGAHEE	0.001
1	SASLYEQHMG	0.001
7	QHMGAHEELK	0.000
4	LYEQHMGAHE	0.000

Table VIII-V4-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
1	NIIHSLMKSI	3.299
2	IIHSLMKSIL	2.047
5	SLMKSILWSK	0.951
6	LMKSILWSKA	0.363
10	ILWSKASGRG	0.137
9	SILWSKASGR	0.008
8	KSILWSKASG	0.002
4	HSLMKSILWS	0.001
7	MKSILWSKAS	0.000
14	KASGRGRREE	0.000
3	IHSLMKSILW	0.000
13	SKASGRGRRE	0.000
12	WSKASGRGRR	0.000
11	LWSKASGRGR	0.000

Table IX-V1-A3-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
754	SLYRNILEK	300.000
417	YLNGNHLTK	60.000

Table IX-V1-A3-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
407	FMNLTRLQK	40.000
433	GLHNLEYLY	36.000
802	TLMYSRPRK	30.000
43	TMLINCEAK	30.000
342	GLLIHCQER	18.000
799	LMETLMYSR	18.000
613	GLLIMFITI	16.200
429	GMFLGLHNL	13.500
174	SLPPNIFRF	13.500
768	QLGITEYLR	12.000
627	GIVVLVLHR	10.800
150	VIEPSAFSK	9.000
415	KLYLNGNHL	9.000
527	GLQQWIQKL	8.100
436	NLEYLYLEY	8.000
431	FLGLHNLEY	8.000
378	LMKSDLVEY	6.000
529	QQWIQKLSK	6.000
546	CTSPGHLDK	3.000
463	VLYLNNNLL	3.000
439	YLYLEYNAI	3.000
2	KLWIHLFYS	2.700
367	KLILAGNII	2.700
297	STKTTSILK	2.000
6	HLFYSSLLA	2.000
632	VLHRRRRYK	2.000
409	NLTRLQKLY	2.000
611	ILGLLIMFI	1.800
337	VLSPSGLLI	1.800
305	KLPTKAPGL	1.800
390	EMLHLGNNR	1.800
158	KLNRLKVLI	1.800
682	MVSPMVHVY	1.800
616	IMFITIVFC	1.500
659	SMYGHKTTH	1.500
628	IVVLVLHRR	1.350
614	LLIMFITIV	1.350
323	QLPGPYCPI	1.350
610	LILGLLIMF	1.350
729	PLTGSNMKY	1.200
453	GTFNPMPKL	1.012
228	QLKTWLENM	0.900

Table IX-V1-A3-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
450	ILPGTFNPM	0.900
615	LIMFITIVF	0.900
609	VLILGLLIM	0.900
255	GSILSRLKK	0.900
482	VPLTKVNLK	0.900
774	YLRKNIAQL	0.900
164	VLILNDNAI	0.900
655	HLQYSMYGH	0.900
86	GLTNAISIH	0.900
71	LLNNGLTML	0.900
656	LQYSMYGHK	0.900
246	VCNSPPFFK	0.900
798	KLMETLMYS	0.810
730	LTGSNMKYK	0.750
681	HMVSPMVHV	0.675
469	NLLQVLPPH	0.675
312	GLIPYITKP	0.608
295	RMSTKTTSI	0.600
630	VLVLHRRRR	0.600
140	FLQADNNFI	0.600
826	NLHAEPDYL	0.600
391	MLHLGNNRI	0.600
68	QLSLLNNGL	0.600
465	YLNNNLLQV	0.600
574	SMPTQTSYL	0.600
70	SLLNNGLTM	0.600
488	NLKTNQFTH	0.600
664	KTTHHTTER	0.600
116	QLHINHNSL	0.600
17	SLHSQTPVL	0.600
187	HLDLRGNQL	0.600
265	SICPTPPVY	0.600
486	KVNLKTNQF	0.600
470	LLQVLPPHI	0.600
110	GLGLLKQLH	0.600
676	SLYEQHMVS	0.600
214	QLEDNKWAC	0.600
781	QLQPDMEAH	0.450
11	SLLACISLH	0.450
178	NIFRFVPLT	0.450
524	DLVGLQQWI	0.405
20	SQTPVLSSR	0.405

Table IX-V1-A3-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
393	HLGNNRIEV	0.400
551	HLDKKELKA	0.400
351	NIESLSDLR	0.400
457	PMPKLKVLY	0.400
812	LVEQTKNEY	0.400
113	LLKQLHINH	0.400
372	GNIIHSLMK	0.360
604	AVPLSVLIL	0.360
741	NQSTEFLSF	0.360
328	YCPIPCNCK	0.300
287	ATSSINDSR	0.300
738	KTTNQSTEF	0.300
728	SPLTGSNMK	0.300
359	RPPPQNPRK	0.300

Table IX-V3-A3-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
7	HMGAHEELK	20.000
2	SLYEQHMGA	3.000
6	QHMGAHEEL	0.001
8	MGAHEELKL	0.001
5	EQHMGAHEE	0.000
1	ASLYEQHMG	0.000
3	LYEQHMGAH	0.000
4	YEQHMGAHE	0.000

Table IX-V4-A3-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
5	LMKSILWSK	135.000
9	ILWSKASGR	20.000
4	SLMKSILWS	0.180
1	IIHSLMKSI	0.045
3	HSLMKSILW	0.003
8	SILWSKASG	0.003
11	WSKASGRGR	0.002
7	KSILWSKAS	0.001
12	SKASGRGRR	0.001
2	IHSLMKSIL	0.001
6	MKSILWSKA	0.000
13	KASGRGRRE	0.000
14	ASGRGRREE	0.000
10	LWSKASGRG	0.000

Table X-V1-HLA-A3-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	**Subsequence**	**Score**
439	YLYLEYNAIK	300.000
798	KLMETLMYSR	121.500
632	VLHRRRRYKK	60.000
768	QLGITEYLRK	40.000
477	HIFSGVPLTK	30.000
210	ILDLQLEDNK	20.000
481	GVPLTKVNLK	18.000
681	HMVSPMVHVY	18.000
616	IMFITIVFCA	13.500
149	TVIEPSAFSK	13.500
158	KLNRLKVLIL	10.800
425	KLSKGMFLGL	10.800
815	QTKNEYFELK	9.000
609	VLILGLLIMF	9.000
245	VVCNSPPFFK	9.000
614	LLIMFITIVF	9.000
811	VLVEQTKNEY	9.000
377	SLMKSDLVEY	9.000
453	GTFNPMPKLK	7.500
781	QLQPDMEAHY	6.000
655	HLQYSMYGHK	6.000
378	LMKSDLVEYF	6.000
75	GLTMLHTNDF	6.000
106	GAFNGLGLLK	6.000
2	KLWIHLFYSS	5.400
86	GLTNAISIHL	5.400
401	VLEEGSFMNL	5.400
42	GTMLINCEAK	4.500
613	GLLIMFITIV	4.050

Table X-V1-HLA-A3-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
627	GIVVLVLHRR	4.050
525	LVGLQQWIQK	4.000
134	GLENLEFLQA	3.600
433	GLHNLEYLYL	3.600
110	GLGLLKQLHI	3.600
6	HLFYSSLLAC	3.000
470	LLQVLPPHIF	3.000
194	QLQTLPYVGF	3.000
290	SINDSRMSTK	3.000
126	ILKEDTFHGL	2.700
357	DLRPPPQNPR	2.700
796	ELKLMETLMY	2.400
546	CTSPGHLDKK	2.250
803	LMYSRPRKVL	2.250
729	PLTGSNMKYK	2.250
369	ILAGNIIHSL	2.025
123	SLEILKEDTF	2.000
765	ELQQLGITEY	1.800
112	GLLKQLHINH	1.800
367	KLILAGNIIH	1.800
78	MLHTNDFSGL	1.800
488	NLKTNQFTHL	1.800
300	TTSILKLPTK	1.500
659	SMYGHKTTHH	1.500
415	KLYLNGNHLT	1.500
568	GLVNNPSMPT	1.350
473	VLPPHIFSGV	1.350
70	SLLNNGLTML	1.350
417	YLNGNHLTKL	1.350
528	LQQWIQKLSK	1.200
409	NLTRLQKLYL	1.200
197	TLPYVGFLEH	1.200
94	HLGFNNIADI	0.900
407	FMNLTRLQKL	0.900
166	ILNDNAIESL	0.900
393	HLGNNRIEVL	0.900
465	YLNNNLLQVL	0.900
469	NLLQVLPPHI	0.900
682	MVSPMVHVYR	0.900
431	FLGLHNLEYL	0.900
337	VLSPSGLLIH	0.900

Table X-V1-HLA-A3-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
232	WLENMPPQSI	0.900
767	QQLGITEYLR	0.810
382	DLVEYFTLEM	0.810
200	YVGFLEHIGR	0.800
611	ILGLLIMFIT	0.675
45	LINCEAKGIK	0.600
600	SLTDAVPLSV	0.600
182	FVPLTHLDLR	0.600
574	SMPTQTSYLM	0.600
647	QMRDNSPVHL	0.600
295	RMSTKTTSIL	0.600
310	APGLIPYITK	0.600
282	SLHLAATSSI	0.600
422	HLTKLSKGMF	0.600
721	LLEQENHSPL	0.600
746	FLSFQDASSL	0.600
630	VLVLHRRRRY	0.600
257	ILSRLKKESI	0.600
336	KVLSPSGLLI	0.540
305	KLPTKAPGLI	0.540
801	ETLMYSRPRK	0.450
753	SSLYRNILEK	0.450
551	HLDKKELKAL	0.450
44	MLINCEAKGI	0.450
441	YLEYNAIKEI	0.450
189	DLRGNQLQTL	0.405
610	LILGLLIMFI	0.405
545	LCTSPGHLDK	0.400
451	LPGTFNPMPK	0.400
71	LLNNGLTMLH	0.400

Table X-V3-HLA-A3-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
8	HMGAHEELKL	1.200
3	SLYEQHMGAH	0.675

Table X-V3-HLA-A3-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
7	QHMGAHEELK	0.045
6	EQHMGAHEEL	0.005
2	ASLYEQHMGA	0.003
1	SASLYEQHMG	0.000
9	MGAHEELKLM	0.000
5	YEQHMGAHEE	0.000
4	LYEQHMGAHE	0.000

Table X-V4-HLA-A3-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
5	SLMKSILWSK	202.500
9	SILWSKASGR	0.600
6	LMKSILWSKA	0.200
1	NIIHSLMKSI	0.068
2	IIHSLMKSIL	0.060
10	ILWSKASGRG	0.030
12	WSKASGRGRR	0.006
4	HSLMKSILWS	0.001
8	KSILWSKASG	0.000
11	LWSKASGRGR	0.000
3	IHSLMKSILW	0.000
14	KASGRGRREE	0.000
7	MKSILWSKAS	0.000
13	SKASGRGRRE	0.000

Table XI-V1-A11-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
529	QQWIQKLSK	2.400
297	STKTTSILK	2.000
546	CTSPGHLDK	2.000

Table XI-V1-A11-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
754	SLYRNILEK	1.600
656	LQYSMYGHK	1.200
150	VIEPSAFSK	1.200
407	FMNLTRLQK	0.800
802	TLMYSRPRK	0.800
417	YLNGNHLTK	0.800
627	GIVVLVLHR	0.720
628	IVVLVLHRR	0.600
440	LYLEYNAIK	0.600
246	VCNSPPFFK	0.600
359	RPPPQNPRK	0.600
664	KTTHHTTER	0.600
43	TMLINCEAK	0.600
730	LTGSNMKYK	0.500
478	IFSGVPLTK	0.400
107	AFNGLGLLK	0.400
372	GNIIHSLMK	0.360
342	GLLIHCQER	0.360
482	VPLTKVNLK	0.300
728	SPLTGSNMK	0.300
420	GNHLTKLSK	0.240
749	FQDASSLYR	0.240
287	ATSSINDSR	0.200
790	YPGAHEELK	0.200
328	YCPIPCNCK	0.200
255	GSILSRLKK	0.180
799	LMETLMYSR	0.160
768	QLGITEYLR	0.160
20	SQTPVLSSR	0.120
454	TFNPMPKLK	0.100
550	GHLDKKELK	0.090
809	RKVLVEQTK	0.090
336	KVLSPSGLL	0.090
462	KVLYLNNNL	0.090
163	KVLILNDNA	0.090
252	FFKGSILSR	0.080
351	NIESLSDLR	0.080
769	LGITEYLRK	0.060
526	VGLQQWIQK	0.060
453	GTFNPMPKL	0.060
42	GTMLINCEA	0.060
629	VVLVLHRRR	0.060

Table XI-V1-A11-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
608	SVLILGLLI	0.060
183	VPLTHLDLR	0.060
486	KVNLKTNQF	0.060
481	GVPLTKVNL	0.060
707	NEKEGSDAK	0.060
291	INDSRMSTK	0.040
182	FVPLTHLDL	0.040
383	LVEYFTLEM	0.040
120	NHNSLEILK	0.040
604	AVPLSVLIL	0.040
633	LHRRRRYKK	0.040
222	CNCDLLQLK	0.040
621	IVFCAAGIV	0.040
46	INCEAKGIK	0.040
632	VLHRRRRYK	0.040
390	EMLHLGNNR	0.036
613	GLLIMFITI	0.036
301	TSILKLPTK	0.030
211	LDLQLEDNK	0.030
738	KTTNQSTEF	0.030
815	QTKNEYFEL	0.030
711	GSDAKHLQR	0.024
433	GLHNLEYLY	0.024
429	GMFLGLHNL	0.024
415	KLYLNGNHL	0.024
816	TKNEYFELK	0.020
155	AFSKLNRLK	0.020
690	YRSPSFGPK	0.020
682	MVSPMVHVY	0.020
87	LTNAISIHL	0.020
601	LTDAVPLSV	0.020
245	VVCNSPPFF	0.020
812	LVEQTKNEY	0.020
547	TSPGHLDKK	0.020
76	LTMLHTNDF	0.020
410	LTRLQKLYL	0.020
698	KHLEEEEER	0.018
367	KLILAGNII	0.018
57	SEISVPPSR	0.018
780	AQLQPDMEA	0.018
701	EEEEERNEK	0.018
615	LIMFITIVF	0.016

Table XI-V1-A11-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
201	VGFLEHIGR	0.016
6	HLFYSSLLA	0.016
591	TNTADTILR	0.016
196	QTLPYVGFL	0.015
148	ITVIEPSAF	0.015
630	VLVLHRRRR	0.012
641	KKQVDEQMR	0.012
86	GLTNAISIH	0.012
527	GLQQWIQKL	0.012
70	SLLNNGLTM	0.012
174	SLPPNIFRF	0.012
488	NLKTNQFTH	0.012
368	LILAGNIIH	0.012

Table XI-V3-A11-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
7	HMGAHEELK	0.400
2	SLYEQHMGA	0.016
3	LYEQHMGAH	0.004
8	MGAHEELKL	0.000
6	QHMGAHEEL	0.000
5	EQHMGAHEE	0.000
4	YEQHMGAHE	0.000
1	ASLYEQHMG	0.000

Table XI-V4-A11-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
5	LMKSILWSK	0.800
9	ILWSKASGR	0.160
12	SKASGRGRR	0.004

1	IIHSLMKSI	0.002
4	SLMKSILWS	0.002
3	HSLMKSILW	0.001
8	SILWSKASG	0.001
11	WSKASGRGR	0.000
6	MKSILWSKA	0.000
2	IHSLMKSIL	0.000
13	KASGRGRRE	0.000
7	KSILWSKAS	0.000
10	LWSKASGRG	0.000
14	ASGRGRREE	0.000

Table XII-V1-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
149	TVIEPSAFSK	9.000
245	VVCNSPPFFK	6.000
42	GTMLINCEAK	6.000
481	GVPLTKVNLK	6.000
525	LVGLQQWIQK	4.000
453	GTFNPMPKLK	3.000
106	GAFNGLGLLK	2.400
477	HIFSGVPLTK	1.600
416	LYLNGNHLTK	1.200
528	LQQWIQKLSK	1.200
815	QTKNEYFELK	1.000
300	TTSILKLPTK	1.000
546	CTSPGHLDKK	1.000
798	KLMETLMYSR	0.960
200	YVGFLEHIGR	0.800
406	SFMNLTRLQK	0.800
439	YLYLEYNAIK	0.800
768	QLGITEYLRK	0.800
632	VLHRRRRYKK	0.800
801	ETLMYSRPRK	0.450
310	APGLIPYITK	0.400
789	HYPGAHEELK	0.400
655	HLQYSMYGHK	0.400
451	LPGTFNPMPK	0.400
689	VYRSPSFGPK	0.400
545	LCTSPGHLDK	0.400
210	ILDLQLEDNK	0.400
590	TTNTADTILR	0.400
290	SINDSRMSTK	0.400

Table XII-V1-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
45	LINCEAKGIK	0.400
682	MVSPMVHVYR	0.400
182	FVPLTHLDLR	0.400
767	QQLGITEYLR	0.360
627	GIVVLVLHRR	0.360
631	LVLHRRRRYK	0.300
221	ACNCDLLQLK	0.200
336	KVLSPSGLLI	0.180
706	RNEKEGSDAK	0.120
254	KGSILSRLKK	0.120
462	KVLYLNNNLL	0.090
163	KVLILNDNAI	0.090
621	IVFCAAGIVV	0.080
748	SFQDASSLYR	0.080
119	INHNSLEILK	0.080
753	SSLYRNILEK	0.060
60	SVPPSRPFQL	0.060
490	KTNQFTHLPV	0.060
700	LEEEEERNEK	0.060
628	IVVLVLHRRR	0.060
608	SVLILGLLIM	0.060
629	VVLVLHRRRR	0.060
296	MSTKTTSILK	0.040
755	LYRNILEKER	0.040
327	PYCPIPCNCK	0.040
154	SAFSKLNRLK	0.040
286	AATSSINDSR	0.040
371	AGNIIHSLMK	0.040
419	NGNHLTKLSK	0.040
350	RNIESLSDLR	0.036
112	GLLKQLHINH	0.036
367	KLILAGNIIH	0.036
738	KTTNQSTEFL	0.030
115	KQLHINHNSL	0.027
433	GLHNLEYLYL	0.024
52	GIKMVSEISV	0.024
110	GLGLLKQLHI	0.024
172	IESLPPNIFR	0.024
158	KLNRLKVLIL	0.024
134	GLENLEFLQA	0.024
616	IMFITIVFCA	0.024

Table XII-V1-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
425	KLSKGMFLGL	0.024
357	DLRPPPQNPR	0.024
86	GLTNAISIHL	0.024
152	EPSAFSKLNR	0.024
389	LEMLHLGNNR	0.024
297	STKTTSILKL	0.020
812	LVEQTKNEYF	0.020
727	HSPLTGSNMK	0.020
686	MVHVYRSPSF	0.020
383	LVEYFTLEML	0.020
358	LRPPPQNPRK	0.020
31	CDSLCNCEEK	0.020
729	PLTGSNMKYK	0.020
423	LTKLSKGMFL	0.020
613	GLLIMFITIV	0.018
181	RFVPLTHLDL	0.018
251	PFFKGSILSR	0.016
178	NIFRFVPLTH	0.016
619	ITIVFCAAGI	0.015
626	AGIVVLVLHR	0.012
640	KKKQVDEQMR	0.012
141	LQADNNFITV	0.012
688	HVYRSPSFGP	0.012
75	GLTMLHTNDF	0.012
609	VLILGLLIMF	0.012
464	LYLNNNLLQV	0.012
614	LLIMFITIVF	0.012
96	GFNNIADIEI	0.012
295	RMSTKTTSIL	0.012
610	LILGLLIMFI	0.012

Table XII-V3-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
7	QHMGAHEELK	0.040
3	SLYEQHMGAH	0.008

Table XII-V3-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
8	HMGAHEELKL	0.008
6	EQHMGAHEEL	0.002
2	ASLYEQHMGA	0.001
4	LYEQHMGAHE	0.000
1	SASLYEQHMG	0.000
9	MGAHEELKLM	0.000
5	YEQHMGAHEE	0.000

Table XII-V4-HLA-A11-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
5	SLMKSILWSK	1.600
9	SILWSKASGR	0.120
2	IIHSLMKSIL	0.004
12	WSKASGRGRR	0.004
6	LMKSILWSKA	0.004
1	NIIHSLMKSI	0.003
10	ILWSKASGRG	0.001
11	LWSKASGRGR	0.000
3	IHSLMKSILW	0.000
8	KSILWSKASG	0.000
4	HSLMKSILWS	0.000
14	KASGRGRREE	0.000
7	MKSILWSKAS	0.000
13	SKASGRGRRE	0.000

Table XIII-V1-HLA-A24-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
443	EYNAIKEIL	420.000
789	HYPGAHEEL	330.000
819	EYFELKANL	288.000

Table XIII-V1-HLA-A24-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
804	MYSRPRKVL	200.000
8	FYSSLLACI	60.000
386	YFTLEMLHL	20.000
139	EFLQADNNF	18.000
462	KVLYLNNNL	17.280
350	RNIESLSDL	14.400
599	RSLTDAVPL	12.000
336	KVLSPSGLL	12.000
305	KLPTKAPGL	12.000
736	KYKTTNQST	12.000
580	SYLMVTTPA	10.500
415	KLYLNGNHL	9.600
272	VYEEHEDPS	9.000
202	GFLEHIGRI	9.000
438	EYLYLEYNA	9.000
466	LNNNLLQVL	8.640
767	QQLGITEYL	8.400
203	FLEHIGRIL	8.400
607	LSVLILGLL	8.400
87	LTNAISIHL	8.400
537	KNTVTDDIL	8.000
219	KWACNCDLL	8.000
758	NILEKEREL	7.920
408	MNLTRLQKL	7.920
527	GLQQWIQKL	7.920
416	LYLNGNHLT	7.500
199	PYVGFLEHI	7.500
486	KVNLKTNQF	7.200
109	NGLGLLKQL	7.200
196	QTLPYVGFL	7.200
133	HGLENLEFL	7.200
225	DLLQLKTWL	7.200
83	DFSGLTNAI	7.200
456	NPMPKLKVL	7.200
561	NSEILCPGL	7.200
501	NILDDLDLL	7.200
500	SNILDDLDL	6.000
221	ACNCDLLQL	6.000
71	LLNNGLTML	6.000
604	AVPLSVLIL	6.000
182	FVPLTHLDL	6.000

Table XIII-V1-HLA-A24-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
347	CQERNIESL	6.000
669	TTERPSASL	6.000
10	SSLLACISL	6.000
590	TTNTADTIL	6.000
481	GVPLTKVNL	6.000
432	LGLHNLEYL	6.000
61	VPPSRPFQL	6.000
394	LGNNRIEVL	6.000
574	SMPTQTSYL	6.000
739	TTNQSTEFL	6.000
68	QLSLLNNGL	5.760
625	AAGIVVLVL	5.600
370	LAGNIIHSL	5.600
593	TADTILRSL	5.600
657	QYSMYGHKT	5.500
154	SAFSKLNRL	4.800
517	NPWDCSCDL	4.800
463	VLYLNNNLL	4.800
752	ASSLYRNIL	4.800
207	IGRILDLQL	4.800
713	DAKHLQRSL	4.800
116	QLHINHNSL	4.800
187	HLDLRGNQL	4.800
426	LSKGMFLGL	4.800
453	GTFNPMPKL	4.400
815	QTKNEYFEL	4.400
418	LNGNHLTKL	4.400
738	KTTNQSTEF	4.400
615	LIMFITIVF	4.200
89	NAISIHLGF	4.200
4	WIHLFYSSL	4.000
26	SSRGSCDSL	4.000
106	GAFNGLGLL	4.000
826	NLHAEPDYL	4.000
429	GMFLGLHNL	4.000
544	ILCTSPGHL	4.000
458	MPKLKVLYL	4.000
159	LNRLKVLIL	4.000
692	SPSFGPKHL	4.000
623	FCAAGIVVL	4.000
296	MSTKTTSIL	4.000

Table XIII-V1-HLA-A24-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
17	SLHSQTPVL	4.000
747	LSFQDASSL	4.000
316	YITKPSTQL	4.000
119	INHNSLEIL	4.000
520	DCSCDLVGL	4.000
405	GSFMNLTRL	4.000
105	IGAFNGLGL	4.000
774	YLRKNIAQL	4.000
410	LTRLQKLYL	4.000
167	LNDNAIESL	4.000
130	DTFHGLENL	4.000
309	KAPGLIPYI	3.600
158	KLNRLKVLI	3.600
76	LTMLHTNDF	3.600
59	ISVPPSRPF	3.600

Table XIII-V3-HLA-A24-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
8	MGAHEELKL	4.400
3	LYEQHMGAH	0.750
6	QHMGAHEEL	0.660
2	SLYEQHMGA	0.120
1	ASLYEQHMG	0.015
5	EQHMGAHEE	0.011
7	HMGAHEELK	0.010
4	YEQHMGAHE	0.002

Table XIII-V4-HLA-A24-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Start	**Subsequence**	**Score**
1	IIHSLMKSI	1.200
2	IHSLMKSIL	0.400
7	KSILWSKAS	0.300
4	SLMKSILWS	0.150
3	HSLMKSILW	0.150
13	KASGRGRRE	0.020
8	SILWSKASG	0.015
5	LMKSILWSK	0.014
6	MKSILWSKA	0.013
14	ASGRGRREE	0.011
10	LWSKASGRG	0.010
11	WSKASGRGR	0.010
9	ILWSKASGR	0.010
12	SKASGRGRR	0.001

Table XIV-V1-HLA-A24-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	**Subsequence**	**Score**
773	EYLRKNIAQL	300.000
385	EYFTLEMLHL	200.000
438	EYLYLEYNAI	90.000
181	RFVPLTHLDL	72.000
202	GFLEHIGRIL	50.400
677	LYEQHMVSPM	37.500
315	PYITKPSTQL	30.000
252	FFKGSILSRL	28.000
622	VFCAAGIVVL	20.000
179	IFRFVPLTHL	20.000
359	RPPPQNPRKL	15.840
462	KVLYLNNNLL	14.400
115	KQLHINHNSL	14.400
757	RNILEKEREL	13.200
832	DYLEVLEQQT	12.960
691	RSPSFGPKHL	12.000
428	KGMFLGLHNL	12.000
158	KLNRLKVLIL	12.000
131	TFHGLENLEF	11.000
425	KLSKGMFLGL	9.600
150	VIEPSAFSKL	9.504
139	EFLQADNNFI	9.000
102	DIEIGAFNGL	8.640
465	YLNNNLLQVL	8.640
67	FQLSLLNNGL	8.640

Table XIV-V1-HLA-A24-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Start	**Subsequence**	**Score**
401	VLEEGSFMNL	8.640
497	LPVSNILDDL	8.400
766	LQQLGITEYL	8.400
96	GFNNIADIEI	8.250
738	KTTNQSTEFL	8.000
380	KSDLVEYFTL	8.000
295	RMSTKTTSIL	8.000
526	VGLQQWIQKL	7.920
407	FMNLTRLQKL	7.920
580	SYLMVTTPAT	7.500
464	LYLNNNLLQV	7.500
828	HAEPDYLEVL	7.200
329	CPIPCNCKVL	7.200
36	NCEEKDGTML	7.200
346	HCQERNIESL	7.200
166	ILNDNAIESL	7.200
60	SVPPSRPFQL	7.200
605	VPLSVLILGL	7.200
480	SGVPLTKVNL	7.200
603	DAVPLSVLIL	7.200
494	FTHLPVSNIL	6.720
592	NTADTILRSL	6.720
417	YLNGNHLTKL	6.600
118	HINHNSLEIL	6.000
500	SNILDDLDLL	6.000
455	FNPMPKLKVL	6.000
70	SLLNNGLTML	6.000
16	ISLHSQTPVL	6.000
8	FYSSLLACIS	6.000
543	DILCTSPGHL	6.000
249	SPPFFKGSIL	6.000
3	LWIHLFYSSL	6.000
825	ANLHAEPDYL	6.000
398	RIEVLEEGSF	6.000
499	VSNILDDLDL	6.000
721	LLEQENHSPL	6.000
383	LVEYFTLEML	6.000
7	LFYSSLLACI	6.000
516	DNPWDCSCDL	6.000
560	LNSEILCPGL	5.760
126	ILKEDTFHGL	5.760

Table XIV-V1-HLA-A24-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
624	CAAGIVVLVL	5.600
86	GLTNAISIHL	5.600
369	ILAGNIIHSL	5.600
657	QYSMYGHKTT	5.000
804	MYSRPRKVLV	5.000
660	MYGHKTTHHT	5.000
493	QFTHLPVSNI	5.000
647	QMRDNSPVHL	4.800
206	HIGRILDLQL	4.800
488	NLKTNQFTHL	4.800
108	FNGLGLLKQL	4.800
668	HTTERPSASL	4.800
189	DLRGNQLQTL	4.800
78	MLHTNDFSGL	4.800
751	DASSLYRNIL	4.800
548	SPGHLDKKEL	4.400
790	YPGAHEELKL	4.400
297	STKTTSILKL	4.400
814	EQTKNEYFEL	4.400
614	LLIMFITIVF	4.200
217	DNKWACNCDL	4.000
9	YSSLLACISL	4.000
409	NLTRLQKLYL	4.000
713	DAKHLQRSLL	4.000
105	IGAFNGLGLL	4.000
431	FLGLHNLEYL	4.000
433	GLHNLEYLYL	4.000
551	HLDKKELKAL	4.000
556	ELKALNSEIL	4.000
374	IIHSLMKSDL	4.000
601	LTDAVPLSVL	4.000
104	EIGAFNGLGL	4.000
393	HLGNNRIEVL	4.000
404	EGSFMNLTRL	4.000

Table XIV-V3-HLA-A24-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
8	HMGAHEELKL	4.400
6	EQHMGAHEEL	4.400
4	LYEQHMGAHE	0.750
9	MGAHEELKLM	0.500
2	ASLYEQHMGA	0.150
3	SLYEQHMGAH	0.012
1	SASLYEQHMG	0.010
5	YEQHMGAHEE	0.002
7	QHMGAHEELK	0.002

Table XIV-V4-HLA-A24-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
2	IIHSLMKSIL	4.000
1	NIIHSLMKSI	1.800
4	HSLMKSILWS	0.150
6	LMKSILWSKA	0.132
8	KSILWSKASG	0.030
14	KASGRGRREE	0.022
5	SLMKSILWSK	0.021
9	SILWSKASGR	0.015
10	ILWSKASGRG	0.010
3	IHSLMKSILW	0.010
7	MKSILWSKAS	0.010
12	WSKASGRGRR	0.010
11	LWSKASGRGR	0.010
13	SKASGRGRRE	0.001

Table XV-V1-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
456	NPMPKLKVL	240.000

Table XV-V1-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
458	MPKLKVLYL	80.000
692	SPSFGPKHL	80.000
61	VPPSRPFQL	80.000
517	NPWDCSCDL	80.000
604	AVPLSVLIL	60.000
26	SSRGSCDSL	40.000
207	IGRILDLQL	40.000
410	LTRLQKLYL	40.000
159	LNRLKVLIL	40.000
774	YLRKNIAQL	40.000
625	AAGIVVLVL	36.000
336	KVLSPSGLL	30.000
481	GVPLTKVNL	20.000
182	FVPLTHLDL	20.000
462	KVLYLNNNL	20.000
652	SPVHLQYSM	20.000
575	MPTQTSYLM	20.000
752	ASSLYRNIL	18.000
370	LAGNIIHSL	12.000
154	SAFSKLNRL	12.000
713	DAKHLQRSL	12.000
221	ACNCDLLQL	12.000
106	GAFNGLGLL	12.000
249	SPPFFKGSI	8.000
306	LPTKAPGLI	8.000
250	PPFFKGSIL	8.000
360	PPPQNPRKL	8.000
453	GTFNPMPKL	6.000
310	APGLIPYIT	6.000
316	YITKPSTQL	6.000
400	EVLEEGSFM	5.000
429	GMFLGLHNL	4.000
418	LNGNHLTKL	4.000
544	ILCTSPGHL	4.000
826	NLHAEPDYL	4.000
350	RNIESLSDL	4.000
4	WIHLFYSSL	4.000
501	NILDDLDLL	4.000
109	NGLGLLKQL	4.000
607	LSVLILGLL	4.000
71	LLNNGLTML	4.000

Table XV-V1-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
599	RSLTDAVPL	4.000
739	TTNQSTEFL	4.000
87	LTNAISIHL	4.000
130	DTFHGLENL	4.000
415	KLYLNGNHL	4.000
175	LPPNIFRFV	4.000
105	IGAFNGLGL	4.000
296	MSTKTTSIL	4.000
63	PSRPFQLSL	4.000
590	TTNTADTIL	4.000
767	QQLGITEYL	4.000
133	HGLENLEFL	4.000
500	SNILDDLDL	4.000
305	KLPTKAPGL	4.000
394	LGNNRIEVL	4.000
815	QTKNEYFEL	4.000
466	LNNNLLQVL	4.000
520	DCSCDLVGL	4.000
747	LSFQDASSL	4.000
623	FCAAGIVVL	4.000
574	SMPTQTSYL	4.000
527	GLQQWIQKL	4.000
426	LSKGMFLGL	4.000
329	CPIPCNCKV	4.000
474	LPPHIFSGV	4.000
10	SSLLACISL	4.000
68	QLSLLNNGL	4.000
405	GSFMNLTRL	4.000
758	NILEKEREL	4.000
17	SLHSQTPVL	4.000
225	DLLQLKTWL	4.000
119	INHNSLEIL	4.000
408	MNLTRLQKL	4.000
463	VLYLNNNLL	4.000
537	KNTVTDDIL	4.000
116	QLHINHNSL	4.000
196	QTLPYVGFL	4.000
432	LGLHNLEYL	4.000
258	LSRLKKESI	4.000
593	TADTILRSL	3.600
792	GAHEELKLM	3.000

Table XV-V1-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
674	SASLYEQHM	3.000
371	AGNIIHSLM	3.000
597	ILRSLTDAV	2.000
608	SVLILGLLI	2.000
807	RPRKVLVEQ	2.000
805	YSRPRKVLV	2.000
498	PVSNILDDL	2.000
364	NPRKLILAG	2.000
339	SPSGLLIHC	2.000
586	TPATTTNTA	2.000
278	DPSGSLHLA	2.000
314	IPYITKPST	2.000
714	AKHLQRSLL	1.800
361	PPQNPRKLI	1.800
669	TTERPSASL	1.800
234	ENMPPQSII	1.800
383	LVEYFTLEM	1.500

Table XV-V3-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
8	MGAHEELKL	4.000
6	QHMGAHEEL	1.200
2	SLYEQHMGA	0.100
1	ASLYEQHMG	0.030
5	EQHMGAHEE	0.010
7	HMGAHEELK	0.010
4	YEQHMGAHE	0.001
3	LYEQHMGAH	0.000

Table XV-V4-HLA-B7-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
1	IIHSLMKSI	0.400
2	IHSLMKSIL	0.400
4	SLMKSILWS	0.060
14	ASGRGRREE	0.045
13	KASGRGRRE	0.030
3	HSLMKSILW	0.020
7	KSILWSKAS	0.020
8	SILWSKASG	0.010
11	WSKASGRGR	0.010
9	ILWSKASGR	0.010
6	MKSILWSKA	0.010
5	LMKSILWSK	0.010
12	SKASGRGRR	0.002
10	LWSKASGRG	0.001

Table XVI-V1-HLA-B7-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
249	SPPFFKGSIL	80.000
548	SPGHLDKKEL	80.000
497	LPVSNILDDL	80.000
329	CPIPCNCKVL	80.000
790	YPGAHEELKL	80.000
605	VPLSVLILGL	80.000
359	RPPPQNPRKL	80.000
189	DLRGNQLQTL	40.000
647	QMRDNSPVHL	40.000
566	CPGLVNNPSM	20.000
807	RPRKVLVEQT	20.000
462	KVLYLNNNLL	20.000
60	SVPPSRPFQL	20.000
713	DAKHLQRSLL	18.000
751	DASSLYRNIL	18.000
603	DAVPLSVLIL	12.000
624	CAAGIVVLVL	12.000
428	KGMFLGLHNL	12.000
825	ANLHAEPDYL	12.000

Table XVI-V1-HLA-B7-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
220	WACNCDLLQL	12.000
803	LMYSRPRKVL	9.000
198	LPYVGFLEHI	8.000
361	PPQNPRKLIL	8.000
176	PPNIFRFVPL	8.000
475	PPHIFSGVPL	8.000
62	PPSRPFQLSL	8.000
179	IFRFVPLTHL	6.000
668	HTTERPSASL	6.000
383	LVEYFTLEML	6.000
608	SVLILGLLIM	5.000
393	HLGNNRIEVL	4.000
589	TTTNTADTIL	4.000
738	KTTNQSTEFL	4.000
78	MLHTNDFSGL	4.000
16	ISLHSQTPVL	4.000
9	YSSLLACISL	4.000
814	EQTKNEYFEL	4.000
407	FMNLTRLQKL	4.000
575	MPTQTSYLMV	4.000
4	WIHLFYSSLL	4.000
417	YLNGNHLTKL	4.000
63	PSRPFQLSLL	4.000
757	RNILEKEREL	4.000
108	FNGLGLLKQL	4.000
409	NLTRLQKLYL	4.000
556	ELKALNSEIL	4.000
166	ILNDNAIESL	4.000
217	DNKWACNCDL	4.000
364	NPRKLILAGN	4.000
295	RMSTKTTSIL	4.000
517	NPWDCSCDLV	4.000
499	VSNILDDLDL	4.000
465	YLNNNLLQVL	4.000
104	EIGAFNGLGL	4.000
346	HCQERNIESL	4.000
691	RSPSFGPKHL	4.000
433	GLHNLEYLYL	4.000
126	ILKEDTFHGL	4.000
526	VGLQQWIQKL	4.000
488	NLKTNQFTHL	4.000

Table XVI-V1-HLA-B7-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
297	STKTTSILKL	4.000
115	KQLHINHNSL	4.000
560	LNSEILCPGL	4.000
334	NCKVLSPSGL	4.000
156	FSKLNRLKVL	4.000
195	LQTLPYVGFL	4.000
86	GLTNAISIHL	4.000
592	NTADTILRSL	4.000
431	FLGLHNLEYL	4.000
746	FLSFQDASSL	4.000
423	LTKLSKGMFL	4.000
158	KLNRLKVLIL	4.000
369	ILAGNIIHSL	4.000
206	HIGRILDLQL	4.000
516	DNPWDCSCDL	4.000
494	FTHLPVSNIL	4.000
500	SNILDDLDLL	4.000
404	EGSFMNLTRL	4.000
766	LQQLGITEYL	4.000
455	FNPMPKLKVL	4.000
480	SGVPLTKVNL	4.000
105	IGAFNGLGLL	4.000
25	LSSRGSCDSL	4.000
236	MPPQSIIGDV	4.000
67	FQLSLLNNGL	4.000
374	IIHSLMKSDL	4.000
543	DILCTSPGHL	4.000
70	SLLNNGLTML	4.000
118	HINHNSLEIL	4.000
425	KLSKGMFLGL	4.000
828	HAEPDYLEVL	3.600
287	ATSSINDSRM	3.000
370	LAGNIIHSLM	3.000
22	TPVLSSRGSC	3.000
278	DPSGSLHLAA	2.000
324	LPGPYCPIPC	2.000
482	VPLTKVNLKT	2.000
163	KVLILNDNAI	2.000
326	GPYCPIPCNC	2.000
336	KVLSPSGLLI	2.000

Table XVI-V3-HLA-B7-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
8	HMGAHEELKL	4.000
6	EQHMGAHEEL	4.000
9	MGAHEELKLM	1.000
2	ASLYEQHMGA	0.300
1	SASLYEQHMG	0.030
3	SLYEQHMGAH	0.010
7	QHMGAHEELK	0.003
5	YEQHMGAHEE	0.001
4	LYEQHMGAHE	0.000

Table XVI-V4-HLA-B7-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
2	IIHSLMKSIL	4.000
1	NIIHSLMKSI	0.400
6	LMKSILWSKA	0.100
14	KASGRGRREE	0.045
5	SLMKSILWSK	0.030
4	HSLMKSILWS	0.020
12	WSKASGRGRR	0.015
8	KSILWSKASG	0.010
10	ILWSKASGRG	0.010
9	SILWSKASGR	0.010
7	MKSILWSKAS	0.002
3	IHSLMKSILW	0.002
13	SKASGRGRRE	0.001
11	LWSKASGRGR	0.001

Table XVII-V1-HLA-B35-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
458	MPKLKVLYL	60.000
652	SPVHLQYSM	40.000
575	MPTQTSYLM	40.000
517	NPWDCSCDL	40.000
456	NPMPKLKVL	20.000
692	SPSFGPKHL	20.000
61	VPPSRPFQL	20.000
792	GAHEELKLM	18.000
26	SSRGSCDSL	15.000
426	LSKGMFLGL	15.000
599	RSLTDAVPL	15.000
727	HSPLTGSNM	10.000
288	TSSINDSRM	10.000
713	DAKHLQRSL	9.000
378	LMKSDLVEY	9.000
306	LPTKAPGLI	8.000
249	SPPFFKGSI	8.000
747	LSFQDASSL	7.500
228	QLKTWLENM	6.000
674	SASLYEQHM	6.000
400	EVLEEGSFM	6.000
258	LSRLKKESI	6.000
796	ELKLMETLM	6.000
752	ASSLYRNIL	5.000
607	LSVLILGLL	5.000
10	SSLLACISL	5.000
59	ISVPPSRPF	5.000
296	MSTKTTSIL	5.000
405	GSFMNLTRL	5.000
815	QTKNEYFEL	4.500
350	RNIESLSDL	4.000
329	CPIPCNCKV	4.000
474	LPPHIFSGV	4.000
782	LQPDMEAHY	4.000
65	RPFQLSLLN	4.000
175	LPPNIFRFV	4.000
805	YSRPRKVLV	3.000
774	YLRKNIAQL	3.000
154	SAFSKLNRL	3.000
410	LTRLQKLYL	3.000
207	IGRILDLQL	3.000
370	LAGNIIHSL	3.000
106	GAFNGLGLL	3.000
156	FSKLNRLKV	3.000
501	NILDDLDLL	3.000
423	LTKLSKGMF	3.000
625	AAGIVVLVL	3.000
159	LNRLKVLIL	3.000

Table XVII-V1-HLA-B35-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
89	NAISIHLGF	3.000
309	KAPGLIPYI	2.400
609	VLILGLLIM	2.000
339	SPSGLLIHC	2.000
450	ILPGTFNPM	2.000
415	KLYLNGNHL	2.000
133	HGLENLEFL	2.000
360	PPPQNPRKL	2.000
278	DPSGSLHLA	2.000
738	KTTNQSTEF	2.000
422	HLTKLSKGM	2.000
586	TPATTTNTA	2.000
314	IPYITKPST	2.000
310	APGLIPYIT	2.000
336	KVLSPSGLL	2.000
778	NIAQLQPDM	2.000
766	LQQLGITEY	2.000
326	GPYCPIPCN	2.000
409	NLTRLQKLY	2.000
631	LVLHRRRRY	2.000
70	SLLNNGLTM	2.000
265	SICPTPPVY	2.000
572	NPSMPTQTS	2.000
462	KVLYLNNNL	2.000
305	KLPTKAPGL	2.000
192	GNQLQTLPY	2.000
825	ANLHAEPDY	2.000
566	CPGLVNNPS	2.000
684	SPMVHVYRS	2.000
250	PPFFKGSIL	2.000
433	GLHNLEYLY	2.000
486	KVNLKTNQF	2.000
331	IPCNCKVLS	2.000
537	KNTVTDDIL	2.000
431	FLGLHNLEY	2.000
758	NILEKEREL	2.000
22	TPVLSSRGS	2.000
152	EPSAFSKLN	2.000
682	MVSPMVHVY	2.000
371	AGNIIHSLM	2.000
650	DNSPVHLQY	2.000

Table XVII-V1-HLA-B35-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
826	NLHAEPDYL	1.500
500	SNILDDLDL	1.500
148	ITVIEPSAF	1.500
520	DCSCDLVGL	1.500
221	ACNCDLLQL	1.500
561	NSEILCPGL	1.500
63	PSRPFQLSL	1.500
293	DSRMSTKTT	1.500
741	NQSTEFLSF	1.500
675	ASLYEQHMV	1.500
100	IADIEIGAF	1.350

Table XVII-V3-HLA-B35-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
8	MGAHEELKL	1.500
2	SLYEQHMGA	0.200
6	QHMGAHEEL	0.100
1	ASLYEQHMG	0.075
5	EQHMGAHEE	0.010
7	HMGAHEELK	0.010
4	YEQHMGAHE	0.001
3	LYEQHMGAH	0.000

Table XVII-V4-HLA-B35-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Start	Subsequence	Score
3	HSLMKSILW	2.500
7	KSILWSKAS	1.000
1	IIHSLMKSI	0.400
11	WSKASGRGR	0.150

2	IHSLMKSIL	0.100
4	SLMKSILWS	0.100
13	KASGRGRRE	0.060
14	ASGRGRREE	0.050
5	LMKSILWSK	0.030
6	MKSILWSKA	0.010
9	ILWSKASGR	0.010
8	SILWSKASG	0.010
10	LWSKASGRG	0.001
12	SKASGRGRR	0.001

Table XVIII-V1-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
319	KPSTQLPGPY	80.000
566	CPGLVNNPSM	40.000
728	SPLTGSNMKY	40.000
572	NPSMPTQTSY	40.000
652	SPVHLQYSMY	40.000
359	RPPPQNPRKL	40.000
456	NPMPKLKVLY	40.000
548	SPGHLDKKEL	30.000
790	YPGAHEELKL	30.000
329	CPIPCNCKVL	20.000
249	SPPFFKGSIL	20.000
605	VPLSVLILGL	20.000
497	LPVSNILDDL	20.000
156	FSKLNRLKVL	15.000
824	KANLHAEPDY	12.000
807	RPRKVLVEQT	12.000
747	LSFQDASSLY	10.000
691	RSPSFGPKHL	10.000
264	ESICPTPPVY	10.000
651	NSPVHLQYSM	10.000
69	LSLLNNGLTM	10.000
796	ELKLMETLMY	9.000
713	DAKHLQRSLL	9.000
198	LPYVGFLEHI	8.000
517	NPWDCSCDLV	8.000
499	VSNILDDLDL	7.500
126	ILKEDTFHGL	6.000
370	LAGNIIHSLM	6.000
458	MPKLKVLYLN	6.000

Table XVIII-V1-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
364	NPRKLILAGN	6.000
647	QMRDNSPVHL	6.000
446	AIKEILPGTF	6.000
535	LSKNTVTDDI	6.000
25	LSSRGSCDSL	5.000
9	YSSLLACISL	5.000
173	ESLPPNIFRF	5.000
16	ISLHSQTPVL	5.000
380	KSDLVEYFTL	4.500
220	WACNCDLLQL	4.500
435	HNLEYLYLEY	4.000
236	MPPQSIIGDV	4.000
382	DLVEYFTLEM	4.000
35	CNCEEKDGTM	4.000
575	MPTQTSYLMV	4.000
777	KNIAQLQPDM	4.000
191	RGNQLQTLPY	4.000
65	RPFQLSLLNN	4.000
811	VLVEQTKNEY	4.000
46	INCEAKGIKM	4.000
556	ELKALNSEIL	3.000
99	NIADIEIGAF	3.000
378	LMKSDLVEYF	3.000
751	DASSLYRNIL	3.000
423	LTKLSKGMFL	3.000
488	NLKTNQFTHL	3.000
377	SLMKSDLVEY	3.000
334	NCKVLSPSGL	3.000
603	DAVPLSVLIL	3.000
624	CAAGIVVLVL	3.000
217	DNKWACNCDL	3.000
297	STKTTSILKL	3.000
189	DLRGNQLQTL	3.000
170	NAIESLPPNI	2.400
475	PPHIFSGVPL	2.000
607	LSVLILGLLI	2.000
346	HCQERNIESL	2.000
295	RMSTKTTSIL	2.000
166	ILNDNAIESL	2.000
630	VLVLHRRRRY	2.000
765	ELQQLGITEY	2.000

Table XVIII-V1-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
115	KQLHINHNSL	2.000
61	VPPSRPFQLS	2.000
278	DPSGSLHLAA	2.000
432	LGLHNLEYLY	2.000
757	RNILEKEREL	2.000
227	LQLKTWLENM	2.000
91	ISIHLGFNNI	2.000
738	KTTNQSTEFL	2.000
176	PPNIFRFVPL	2.000
781	QLQPDMEAHY	2.000
592	NTADTILRSL	2.000
158	KLNRLKVLIL	2.000
84	FSGLTNAISI	2.000
668	HTTERPSASL	2.000
248	NSPPFFKGSI	2.000
287	ATSSINDSRM	2.000
428	KGMFLGLHNL	2.000
681	HMVSPMVHVY	2.000
22	TPVLSSRGSC	2.000
449	EILPGTFNPM	2.000
425	KLSKGMFLGL	2.000
408	MNLTRLQKLY	2.000
560	LNSEILCPGL	2.000
361	PPQNPRKLIL	2.000
62	PPSRPFQLSL	2.000
574	SMPTQTSYLM	2.000
482	VPLTKVNLKT	2.000
324	LPGPYCPIPC	2.000
462	KVLYLNNNLL	2.000
608	SVLILGLLIM	2.000

Table XVIII-V3-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
9	MGAHEELKLM	3.000
8	HMGAHEELKL	1.500

Table XVIII-V3-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
6	EQHMGAHEEL	1.000
2	ASLYEQHMGA	0.500
1	SASLYEQHMG	0.045
3	SLYEQHMGAH	0.020
5	YEQHMGAHEE	0.001
7	QHMGAHEELK	0.001
4	LYEQHMGAHE	0.000

Table XVIII-V4-HLA-B35-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Start	Subsequence	Score
2	IIHSLMKSIL	1.000
4	HSLMKSILWS	0.500
1	NIIHSLMKSI	0.400
6	LMKSILWSKA	0.300
12	WSKASGRGRR	0.150
8	KSILWSKASG	0.100
14	KASGRGRREE	0.060
3	IHSLMKSILW	0.050
7	MKSILWSKAS	0.010
9	SILWSKASGR	0.010
10	ILWSKASGRG	0.010
5	SLMKSILWSK	0.010
13	SKASGRGRRE	0.001
11	LWSKASGRGR	0.001

Table V – 158P1D7 v.6 – HLA A1-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score

Table V – 158P1D7 v.6 – HLA A1-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
7	LMNPSFGPK	1.000
5	HSLMNPSFG	0.015
1	GNIIHSLMN	0.013
4	IHSLMNPSF	0.010
3	IIHSLMNPS	0.010
8	MNPSFGPKH	0.005
6	SLMNPSFGP	0.005
2	NIIHSLMNP	0.005
15	KHLEEEEER	0.005
9	NPSFGPKHL	0.003
11	SFGPKHLEE	0.003
10	PSFGPKHLE	0.000
12	FGPKHLEEE	0.000
13	GPKHLEEEE	0.000
14	PKHLEEEEE	0.000

Table VI – 158P1D7 v.6 – HLA A1-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
4	IIHSIMNPSF	0.200
7	SLMNpSFGPK	0.200
8	LMNPsFGPKH	0.100
1	AGNIiHSLMN	0.013
3	NIIHsLMNPS	0.010
6	HSLMnPSFGP	0.007
9	MNPSfGPKHL	0.003
2	GNIIhSLMNP	0.001
11	PSFGpKHLEE	0.001
15	PKHLeEEEER	0.001
5	IHSLmNPSFG	0.001
12	SFGPkHLEEE	0.001

Table VI – 158P1D7 v.6 – HLA A1-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
10	NPSFgPKHLE	0.000
13	FGPKhLEEEE	0.000
14	GPKHIEEEEE	0.000

Table VII – 158P1D7 v.6 – HLA A0201-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
6	SLMNPSFGP	0.320
9	NPSFGPKHL	0.139
3	IIHSLMNPS	0.040
2	NIIHSLMNP	0.005
7	LMNPSFGPK	0.005
8	MNPSFGPKH	0.003
12	FGPKHLEEE	0.001
1	GNIIHSLMN	0.000
5	HSLMNPSFG	0.000
15	KHLEEEEER	0.000
4	IHSLMNPSF	0.000
11	SFGPKHLEE	0.000
10	PSFGPKHLE	0.000
13	GPKHLEEEE	0.000
14	PKHLEEEEE	0.000

Table VIII – 158P1D7 v.6 – HLA A0201-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
8	LMNPsFGPKH	0.348
9	MNPSfGPKHL	0.237
3	NIIHsLMNPS	0.024
4	IIHSlMNPSF	0.017
7	SLMNpSFGPK	0.014
1	AGNIiHSLMN	0.000
5	IHSLmNPSFG	0.000
2	GNIIhSLMNP	0.000
13	FGPKhLEEEE	0.000
10	NPSFgPKHLE	0.000
6	HSLMnPSFGP	0.000
12	SFGPkHLEEE	0.000
11	PSFGpKHLEE	0.000
14	GPKHIEEEEE	0.000
15	PKHLeEEEER	0.000

Table IX – 158P1D7 v.6 – HLA A3-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
7	LMNPSFGPK	27.000
6	SLMNPSFGP	0.135
15	KHLEEEEER	0.027
2	NIIHSLMNP	0.009
3	IIHSLMNPS	0.006
9	NPSFGPKHL	0.003
4	IHSLMNPSF	0.002
8	MNPSFGPKH	0.001
13	GPKHLEEEE	0.001
1	GNIIHSLMN	0.000
5	HSLMNPSFG	0.000
10	PSFGPKHLE	0.000
11	SFGPKHLEE	0.000
12	FGPKHLEEE	0.000
14	PKHLEEEEE	0.000

Table X – 158P1D7 v.6 – HLA A3-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
7	LMNPSFGPK	27.000
6	SLMNPSFGP	0.135
15	KHLEEEEER	0.027
2	NIIHSLMNP	0.009
3	IIHSLMNPS	0.006
9	NPSFGPKHL	0.003
4	IHSLMNPSF	0.002
8	MNPSFGPKH	0.001
13	GPKHLEEEE	0.001
1	GNIIHSLMN	0.000
5	HSLMNPSFG	0.000
10	PSFGPKHLE	0.000
11	SFGPKHLEE	0.000
12	FGPKHLEEE	0.000
14	PKHLEEEEE	0.000

Table XI – 158P1D7 v.6 – HLA A1101-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
7	LMNPSFGPK	0.400
15	KHLEEEEER	0.018
6	SLMNPSFGP	0.002
2	NIIHSLMNP	0.001
9	NPSFGPKHL	0.001
13	GPKHLEEEE	0.001
11	SFGPKHLEE	0.000
8	MNPSFGPKH	0.000
3	IIHSLMNPS	0.000
1	GNIIHSLMN	0.000
4	IHSLMNPSF	0.000
5	HSLMNPSFG	0.000
12	FGPKHLEEE	0.000

Table XI – 158P1D7 v.6 – HLA A1101-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
10	PSFGPKHLE	0.000
14	PKHLEEEEE	0.000

Table XII – 158P1D7 v.6 – HLA A1101-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
7	SLMNpSFGPK	0.800
4	IIHSlMNPSF	0.004
8	LMNPsFGPKH	0.004
3	NIIHsLMNPS	0.001
14	GPKHIEEEEE	0.001
15	PKHLeEEEER	0.000
2	GNIIhSLMNP	0.000
9	MNPSfGPKHL	0.000
10	NPSFgPKHLE	0.000
12	SFGPkHLEEE	0.000
6	HSLMnPSFGP	0.000
1	AGNIiHSLMN	0.000
13	FGPKhLEEEE	0.000
5	IHSLmNPSFG	0.000
11	PSFGpKHLEE	0.000

Table XIII – 158P1D7 v.6 – HLA A24-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
9	NPSFGPKHL	4.000

Table XIII – 158P1D7 v.6 – HLA A24-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
4	IHSLMNPSF	0.200
1	GNIIHSLMN	0.150
3	IIHSLMNPS	0.144
11	SFGPKHLEE	0.066
7	LMNPSFGPK	0.022
12	FGPKHLEEE	0.017
8	MNPSFGPKH	0.017
6	SLMNPSFGP	0.015
5	HSLMNPSFG	0.015
2	NIIHSLMNP	0.015
13	GPKHLEEEE	0.013
15	KHLEEEEER	0.004
10	PSFGPKHLE	0.001
14	PKHLEEEEE	0.000

Table XIV – 158P1D7 v.6 – HLA A24-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
9	MNPSfGPKHL	6.000
4	IIHSlMNPSF	2.000
3	NIIHsLMNPS	0.216
1	AGNIiHSLMN	0.150
12	SFGPkHLEEE	0.066
13	FGPKhLEEEE	0.020
8	LMNPsFGPKH	0.020
7	SLMNpSFGPK	0.018
2	GNIIhSLMNP	0.015
6	HSLMnPSFGP	0.015
14	GPKHlEEEEE	0.011
10	NPSFgPKHLE	0.010
11	PSFGpKHLEE	0.001

Table XIV – 158P1D7 v.6 – HLA A24-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
5	IHSLmNPSFG	0.001
15	PKHLeEEEER	0.000

Table XV – 158P1D7 v.6 – HLA B7-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
9	NPSFGPKHL	80.000
13	GPKHLEEEE	0.200
6	SLMNPSFGP	0.045
3	IIHSLMNPS	0.020
1	GNIIHSLMN	0.020
5	HSLMNPSFG	0.010
7	LMNPSFGPK	0.010
8	MNPSFGPKH	0.010
2	NIIHSLMNP	0.010
12	FGPKHLEEE	0.010
4	IHSLMNPSF	0.002
10	PSFGPKHLE	0.002
15	KHLEEEEER	0.001
11	SFGPKHLEE	0.001
14	PKHLEEEEE	0.000

Table XVI – 158P1D7 v.6 – HLA B7-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score

Table XVI – 158P1D7 v.6 – HLA B7-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
9	MNPSfGPKHL	4.000
10	NPSFgPKHLE	0.300
14	GPKHlEEEEE	0.200
1	AGNIiHSLMN	0.060
7	SLMNpSFGPK	0.030
4	IIHSlMNPSF	0.020
3	NIIHsLMNPS	0.020
6	HSLMnPSFGP	0.015
13	FGPKhLEEEE	0.010
8	LMNPsFGPKH	0.010
2	GNIIhSLMNP	0.010
12	SFGPkHLEEE	0.001
11	PSFGpKHLEE	0.001
5	IHSLmNPSFG	0.001
15	PKHLeEEEER	0.000

Table XVII - 158P1D7 v.6 – HLA B3501-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
9	NPSFGPKHL	20.000
13	GPKHLEEEE	0.600
1	GNIIHSLMN	0.100
4	IHSLMNPSF	0.100
3	IIHSLMNPS	0.100
5	HSLMNPSFG	0.050
7	LMNPSFGPK	0.010
8	MNPSFGPKH	0.010
6	SLMNPSFGP	0.010
2	NIIHSLMNP	0.010
12	FGPKHLEEE	0.010
15	KHLEEEEER	0.006

Table XVII - 158P1D7 v.6 – HLA B3501-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	Subsequence	Score
10	PSFGPKHLE	0.005
11	SFGPKHLEE	0.001
14	PKHLEEEEE	0.000

Table XVIII – 158P1D7 v.6 – HLA B3501-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	Subsequence	Score
9	MNPSfGPKHL	1.000
4	IIHSlMNPSF	1.000
14	GPKHlEEEEE	0.900
10	NPSFgPKHLE	0.200
1	AGNIiHSLMN	0.100
3	NIIHsLMNPS	0.100
6	HSLMnPSFGP	0.050
2	GNIIhSLMNP	0.010
8	LMNPsFGPKH	0.010
13	FGPKhLEEEE	0.010
7	SLMNpSFGPK	0.010
11	PSFGpKHLEE	0.005
12	SFGPkHLEEE	0.001
5	IHSLmNPSFG	0.001
15	PKHLeEEEER	0.000

Table XIX: Motif-bearing Subsequences of the 158P1D7 Protein

Protein Motifs of 158P1D7

N-glycosylation site
Number of matches: 3
1 292-295 NDSR (SEQ ID NO: 45)
2 409-412 NLTR (SEQ ID NO: 46)
3 741-744 NQST (SEQ ID NO: 47)

cAMP- and cGMP-dependent protein kinase phosphorylation site
262-265 KKES (SEQ ID NO: 48)

Protein kinase C phosphorylation site
Number of matches: 3
1 26-28 SSR
2 297-299 STK
3 670-672 TER

Casein kinase II phosphorylation site
Number of matches: 12
1 149-152 TVIE (SEQ ID NO: 49)
2 186-189 THLD (SEQ ID NO: 50)
3 231-234 TWLE (SEQ ID NO: 51)
4 290-293 SIND (SEQ ID NO: 52)
5 354-357 SLSD (SEQ ID NO: 53)
6 510-513 TQID (SEQ ID NO: 54)
7 539-542 TVTD (SEQ ID NO: 55)
8 600-603 SLTD (SEQ ID NO: 56)
9 676-679 SLYE (SEQ ID NO: 57)
10 720-723 SLLE (SEQ ID NO: 58)
11 748-751 SFQD (SEQ ID NO: 59)
12 816-819 TKNE (SEQ ID NO: 60)

Tyrosine kinase phosphorylation site
798-805 KLMETLMY (SEQ ID NO: 61)

N-myristoylation site
Number of matches: 8
1 29-34 GSCDSL (SEQ ID NO: 62)
2 86-91 GLTNAI (SEQ ID NO: 63)
3 106-111 GAFNGL (SEQ ID NO: 64)
4 255-260 GSILSR (SEQ ID NO: 65)
5 405-410 GSFMNL (SEQ ID NO: 66)
6 420-425 GNHLTK (SEQ ID NO: 67)
7 429-434 GMFLGL (SEQ ID NO: 68)
8 481-486 GVPLTK (SEQ ID NO: 69)

Two Protein Motifs were predicted by Pfam
1-Archaeal-ATPase at aa 441-451
2-Leucine rich repeat C-terminal at aa 218-268 and aa 517-567

Table XX: Frequently Occurring Motifs			
Name	avrg. % identity	Description	Potential Function
zf-C2H2	34%	Zinc finger, C2H2 type	Nucleic acid-binding protein functions as transcription factor, nuclear location probable
cytochrome_b_N	68%	Cytochrome b(N-terminal)/b6/petB	membrane bound oxidase, generate superoxide
ig	19%	Immunoglobulin domain	domains are one hundred amino acids long and include a conserved intradomain disulfide bond.
WD40	18%	WD domain, G-beta repeat	tandem repeats of about 40 residues, each containing a Trp-Asp motif. Function in signal transduction and protein interaction
PDZ	23%	PDZ domain	may function in targeting signaling molecules to sub-membranous sites
LRR	28%	Leucine Rich Repeat	short sequence motifs involved in protein-protein interactions
pkinase	23%	Protein kinase domain	conserved catalytic core common to both serine/threonine and tyrosine protein kinases containing an ATP binding site and a catalytic site
PH	16%	PH domain	pleckstrin homology involved in intracellular signaling or as constituents of the cytoskeleton
EGF	34%	EGF-like domain	30-40 amino-acid long found in the extracellular domain of membrane-bound proteins or in secreted proteins
rvt	49%	Reverse transcriptase (RNA-dependent DNA polymerase)	
ank	25%	Ank repeat	Cytoplasmic protein, associates integral membrane proteins to the cytoskeleton
oxidored_q1	32%	NADH-Ubiquinone/plastoquinone (complex I), various chains	membrane associated. Involved in proton translocation across the membrane
efhand	24%	EF hand	calcium-binding domain, consists of a12 residue loop flanked on both sides by a 12 residue alpha-helical domain
rvp	79%	Retroviral aspartyl protease	Aspartyl or acid proteases, centered on a catalytic aspartyl residue

Collagen	42%	Collagen triple helix repeat (20 copies)	extracellular structural proteins involved in formation of connective tissue. The sequence consists of the G-X-Y and the polypeptide chains forms a triple helix.
fn3	20%	Fibronectin type III domain	Located in the extracellular ligand-binding region of receptors and is about 200 amino acid residues long with two pairs of cysteines involved in disulfide bonds
7tm_1	19%	7 transmembrane receptor (rhodopsin family)	seven hydrophobic transmembrane regions, with the N-terminus located extracellularly while the C-terminus is cytoplasmic. Signal through G proteins

Table XXI: TNM CLASSIFICATION OF BLADDER TUMORS

Primary tumor (T)

The suffix (m) should be added to the appropriate T category to indicate multiple tumors. The suffix (is) may be added to any T to indicate the presence of associated carcinoma *in situ*.

- TX Primary tumor cannot be assessed
- TO No evidence of primary tumor
- Ta Noninvasive papillary carcinoma
- Tis Carcinoma *in situ*: "flat tumor"
- T1 Tumor invades sub-epithelial connective tissue
- T2 Tumor invades superficial muscle (inner half)
- T3 Tumor invades deep muscle or perivesical fat
  - T3a Tumor invades deep muscle (outer half)
  - T3b Tumor invades perivesical fat
    - i. microscopically
    - ii. macroscopically (extravesical mass)
- T4 Tumor invades any of the following: prostate, uterus, vagina, pelvic wall, or abdominal wall
  - T4a Tumor invades the prostate, uterus, vagina
  - T4b Tumor invades the pelvic wall or abdominal wall or both

Regional lymph nodes (N)

Regional lymph nodes are those within the true pelvis: all others are distant nodes

- NX Regional lymph nodes cannot be assessed
- N0 No regional lymph node metastasis
- N1 Metastasis in a single lymph node, 2 cm or less in greatest dimension
- N2 Metastasis in a single lymph node, more than 2 cm but not more than 5 cm in greatest dimension, or multiple lymph nodes, none more than 5 cm in greatest dimension
- N3 Metastasis in a lymph node more than 5 cm in greatest dimension

Distant metastasis (M)

- MX Presence of distant metastasis cannot be assessed
- M0 No distant metastasis
- M1 Distant metastasis

Stage grouping

Stage				
	0_a	Ta	N0	M0
	0_{is}	Tis	N0	M0
	I	T1	N0	M0
	II	T2	N0	M0
		T3a	N0	M0
	III	T3b	N0	M0
		T4a	N0	M0
	IV	T4b	N0	M0
		Any T	N1-3	M0
		Any T	Any N	M1

Table XXII-V1-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
436	NLEYLYLEY	32
650	DNSPVHLQY	27
308	TKAPGLIPY	25
812	LVEQTKNEY	25
431	FLGLHNLEY	24
601	LTDAVPLSV	24
192	GNQLQTLPY	23
573	PSMPTQTSY	23
265	SICPTPPVY	22
797	LKLMETLMY	22
1	MKLWIHLFY	21
522	SCDLVGLQQ	21
670	TERPSASLY	21
682	MVSPMVHVY	21
711	GSDAKHLQR	20
729	PLTGSNMKY	20
828	HAEPDYLEV	20
320	PSTQLPGPY	19
441	YLEYNAIKE	19
502	ILDDLDLLT	19
551	HLDKKELKA	19
748	SFQDASSLY	19
223	NCDLLQLKT	18
409	NLTRLQKLY	18
433	GLHNLEYLY	18
546	CTSPGHLDK	18
653	PVHLQYSMY	18
743	STEFLSFQD	18
763	ERELQQLGI	18
793	AHEELKLME	18
817	KNEYFELKA	18
39	EKDGTMLIN	17
47	NCEAKGIKM	17
81	TNDFSGLTN	17
142	QADNNFITV	17
276	HEDPSGSLH	17
388	TLEMLHLGN	17
457	PMPKLKVLY	17
540	VTDDILCTS	17
669	TTERPSASL	17
749	FQDASSLYR	17
766	LQQLGITEY	17
771	ITEYLRKNI	17

Table XXII-V1-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
56	VSEISVPPS	16
380	KSDLVEYFT	16
383	LVEYFTLEM	16
503	LDDLDLLTQ	16
554	KKELKALNS	16
631	LVLHRRRRY	16
825	ANLHAEPDY	16
150	VIEPSAFSK	15
337	VLSPSGLLI	15
378	LMKSDLVEY	15
401	VLEEGSFMN	15
782	LQPDMEAHY	15

Table XXII-V3-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
3	LYEQHMGAH	10
8	MGAHEELKL	8
1	ASLYEQHMG	6
2	SLYEQHMGA	5

Table XXII-V4-HLA-A1-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
3	HSLMKSILW	10
4	SLMKSILWS	9
14	ASGRGRREE	8

Table XXII-V4-HLA-A1-9mers-158P1D7

11	WSKASGRGR	5
12	SKASGRGRR	5
7	KSILWSKAS	4

Table XXIII-V1-HLA-A2-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
71	LLNNGLTML	29
614	LLIMFITIV	29
465	YLNNNLLQV	28
774	YLRKNIAQL	28
429	GMFLGLHNL	27
527	GLQQWIQKL	27
597	ILRSLTDAV	26
17	SLHSQTPVL	25
501	NILDDLDLL	25
611	ILGLLIMFI	25
758	NILEKEREL	25
305	KLPTKAPGL	24
606	PLSVLILGL	24
609	VLILGLLIM	24
624	CAAGIVVLV	24
68	QLSLLNNGL	23
116	QLHINHNSL	23
154	SAFSKLNRL	23
158	KLNRLKVLI	23
164	VLILNDNAI	23
196	QTLPYVGFL	23
370	LAGNIIHSL	23
415	KLYLNGNHL	23
439	YLYLEYNAI	23
463	VLYLNNNLL	23
613	GLLIMFITI	23
803	LMYSRPRKV	23
106	GAFNGLGLL	22
225	DLLQLKTWL	22
312	GLIPYITKP	22
337	VLSPSGLLI	22

Table XXIII-V1-HLA-A2-9mers-158P1D7		
367	KLILAGNII	22
393	HLGNNRIEV	22
470	LLQVLPPHI	22
544	ILCTSPGHL	22
564	ILCPGLVNN	22
574	SMPTQTSYL	22
4	WIHLFYSSL	21
70	SLLNNGLTM	21
92	SIHLGFNNI	21
187	HLDLRGNQL	21
295	RMSTKTTSI	21
309	KAPGLIPYI	21
323	QLPGPYCPI	21
391	MLHLGNNRI	21
446	AIKEILPGT	21
581	YLMVTTPAT	21
604	AVPLSVLIL	21
623	FCAAGIVVL	21
625	AAGIVVLVL	21
681	HMVSPMVHV	21
118	HINHNSLEI	20
130	DTFHGLENL	20
140	FLQADNNFI	20
203	FLEHIGRIL	20
240	SIIGDVVCN	20
316	YITKPSTQL	20
369	ILAGNIIHS	20
453	GTFNPMPKL	20
477	HIFSGVPLT	20
524	DLVGLQQWI	20
593	TADTILRSL	20
754	SLYRNILEK	20
826	NLHAEPDYL	20
45	LINCEAKGI	19
171	AIESLPPNI	19
178	NIFRFVPLT	19
302	SILKLPTKA	19
450	ILPGTFNPM	19
473	VLPPHIFSG	19
502	ILDDLDLLT	19
601	LTDAVPLSV	19
610	LILGLLIMF	19
11	SLLACISLH	18
103	IEIGAFNGL	18

Table XXIII-V1-HLA-A2-9mers-158P1D7		
109	NGLGLLKQL	18
112	GLLKQLHIN	18
133	HGLENLEFL	18
159	LNRLKVLIL	18
167	LNDNAIESL	18
174	SLPPNIFRF	18
190	LRGNQLQTL	18
221	ACNCDLLQL	18
290	SINDSRMST	18
336	KVLSPSGLL	18
344	LIHCQERNI	18
350	RNIESLSDL	18
408	MNLTRLQKL	18
417	YLNGNHLTK	18
418	LNGNHLTKL	18
432	LGLHNLEYL	18
462	KVLYLNNNL	18
466	LNNNLLQVL	18
479	FSGVPLTKV	18
494	FTHLPVSNI	18
551	HLDKKELKA	18
559	ALNSEILCP	18
582	LMVTTPATT	18
596	TILRSLTDA	18
608	SVLILGLLI	18
620	TIVFCAAGI	18
669	TTERPSASL	18
798	KLMETLMYS	18
828	HAEPDYLEV	18
829	AEPDYLEVL	18
48	CEAKGIKMV	17
51	KGIKMVSEI	17
87	LTNAISIHL	17
95	LGFNNIADI	17
157	SKLNRLKVL	17
180	FRFVPLTHL	17
193	NQLQTLPYV	17
202	GFLEHIGRI	17
228	QLKTWLENM	17
256	SILSRLKKE	17
378	LMKSDLVEY	17
394	LGNNRIEVL	17
410	LTRLQKLYL	17
456	NPMPKLKVL	17

Table XXIII-V1-HLA-A2-9mers-158P1D7		
469	NLLQVLPPH	17
481	GVPLTKVNL	17
534	KLSKNTVTD	17
556	ELKALNSEI	17
600	SLTDAVPLS	17
602	TDAVPLSVL	17
616	IMFITIVFC	17
621	IVFCAAGIV	17
716	HLQRSLLEQ	17
720	SLLEQENHS	17
739	TTNQSTEFL	17
770	GITEYLRKN	17
2	KLWIHLFYS	16
8	FYSSLLACI	16
10	SSLLACISL	16
26	SSRGSCDSL	16
44	MLINCEAKG	16
99	NIADIEIGA	16
119	INHNSLEIL	16
123	SLEILKEDT	16
142	QADNNFITV	16
143	ADNNFITVI	16
166	ILNDNAIES	16
182	FVPLTHLDL	16
189	DLRGNQLQT	16
205	EHIGRILDL	16
210	ILDLQLEDN	16
283	LHLAATSSI	16
298	TKTTSILKL	16
329	CPIPCNCKV	16
373	NIIHSLMKS	16
381	SDLVEYFTL	16
405	GSFMNLTRL	16
442	LEYNAIKEI	16
520	DCSCDLVGL	16
603	DAVPLSVLI	16
607	LSVLILGLL	16
767	QQLGITEYL	16
778	NIAQLQPDM	16
805	YSRPRKVLV	16
833	YLEVLEQQT	16
6	HLFYSSLLA	15
12	LLACISLHS	15
53	IKMVSEISV	15

Table XXIII-V1-HLA-A2-9mers-158P1D7		
64	SRPFQLSLL	15
105	IGAFNGLGL	15
126	ILKEDTFHG	15
147	FITVIEPSA	15
161	RLKVLILND	15
209	RILDLQLED	15
226	LLQLKTWLE	15
241	IIGDVVCNS	15
253	FKGSILSRL	15
342	GLLIHCQER	15
347	CQERNIESL	15
354	SLSDLRPPP	15
384	VEYFTLEML	15
426	LSKGMFLGL	15
455	FNPMPKLKV	15
458	MPKLKVLYL	15
495	THLPVSNIL	15
498	PVSNILDDL	15
500	SNILDDLDL	15
504	DDLDLLTQI	15
507	DLLTQIDLE	15
552	LDKKELKAL	15
590	TTNTADTIL	15
627	GIVVLVLHR	15
659	SMYGHKTTH	15
676	SLYEQHMVS	15
713	DAKHLQRSL	15
747	LSFQDASSL	15
815	QTKNEYFEL	15
5	IHLFYSSLL	14
16	ISLHSQTPV	14
33	SLCNCEEKD	14
83	DFSGLTNAI	14
85	SGLTNAISI	14
86	GLTNAISIH	14
90	AISIHLGFN	14
111	LGLLKQLHI	14
127	LKEDTFHGL	14
151	IEPSAFSKL	14
165	LILNDNAIE	14
207	IGRILDLQL	14
233	LENMPPQSI	14
257	ILSRLKKES	14
282	SLHLAATSS	14

Table XXIII-V1-HLA-A2-9mers-158P1D7		
303	ILKLPTKAP	14
330	PIPCNCKVL	14
343	LLIHCQERN	14
368	LILAGNIIH	14
377	SLMKSDLVE	14
383	LVEYFTLEM	14
387	FTLEMLHLG	14
401	VLEEGSFMN	14
422	HLTKLSKGM	14
431	FLGLHNLEY	14
434	LHNLEYLYL	14
506	LDLLTQIDL	14
508	LLTQIDLED	14
532	IQKLSKNTV	14
557	LKALNSEIL	14
562	SEILCPGLV	14
599	RSLTDAVPL	14
675	ASLYEQHMV	14
721	LLEQENHSP	14
722	LEQENHSPL	14
746	FLSFQDASS	14
752	ASSLYRNIL	14
789	HYPGAHEEL	14
792	GAHEELKLM	14
811	VLVEQTKNE	14

Table XXIII-3-HLA-A2-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
2	SLYEQHMGA	20
6	QHMGAHEEL	15
8	MGAHEELKL	15

Table XXIII-4-HLA-A2-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
1	IIHSLMKSI	20
4	SLMKSILWS	18
8	SILWSKASG	16
5	LMKSILWSK	15
9	ILWSKASGR	15
2	IHSLMKSIL	12

Table XXIV-V1-HLA-A0203-9mers-158P1D7		
Pos	123456789	score
NoResultsFound.		

Table XXIV-V3-HLA-A0203-9mers-158P1D7		
Pos	123456789	score
NoResultsFound.		

Table XXIV-V4-HLA-A0203-9mers-158P1D7		
Pos	123456789	score
NoResultsFound.		

Table XXV-V1-HLA-A3-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
754	SLYRNILEK	31
417	YLNGNHLTK	29
150	VIEPSAFSK	26
632	VLHRRRRYK	26
70	SLLNNGLTM	24
265	SICPTPPVY	23

Table XXV-V1-HLA-A3-9mers-158P1D7		
478	IFSGVPLTK	23
682	MVSPMVHVY	23
11	SLLACISLH	22
486	KVNLKTNQF	22
107	AFNGLGLLK	21
189	DLRGNQLQT	21
291	INDSRMSTK	21
415	KLYLNGNHL	21
534	KLSKNTVTD	21
564	ILCPGLVNN	21
631	LVLHRRRRY	21
653	PVHLQYSMY	21
676	SLYEQHMVS	21
688	HVYRSPSFG	21
802	TLMYSRPRK	21
158	KLNRLKVLI	20
367	KLILAGNII	20
431	FLGLHNLEY	20
563	EILCPGLVN	20
608	SVLILGLLI	20
781	QLQPDMEAH	20
809	RKVLVEQTK	20
187	HLDLRGNQL	19
301	TSILKLPTK	19
337	VLSPSGLLI	19
400	EVLEEGSFM	19
409	NLTRLQKLY	19
436	NLEYLYLEY	19
488	NLKTNQFTH	19
609	VLILGLLIM	19
633	LHRRRRYKK	19
729	PLTGSNMKY	19
774	YLRKNIAQL	19
24	VLSSRGSCD	18
86	GLTNAISIH	18
161	RLKVLILND	18
174	SLPPNIFRF	18
179	IFRFVPLTH	18
209	RILDLQLED	18
240	SIIGDVVCN	18
255	GSILSRLKK	18
282	SLHLAATSS	18
368	LILAGNIIH	18
372	GNIIHSLMK	18

Table XXV-V1-HLA-A3-9mers-158P1D7		
377	SLMKSDLVE	18
407	FMNLTRLQK	18
529	QQWIQKLSK	18
546	CTSPGHLDK	18
583	MVTTPATTT	18
628	IVVLVLHRR	18
634	HRRRRYKKK	18
670	TERPSASLY	18
44	MLINCEAKG	17
149	TVIEPSAFS	17
194	QLQTLPYVG	17
305	KLPTKAPGL	17
311	PGLIPYITK	17
312	GLIPYITKP	17
342	GLLIHCQER	17
357	DLRPPPQNP	17
359	RPPPQNPRK	17
412	RLQKLYLNG	17
433	GLHNLEYLY	17
460	KLKVLYLNN	17
465	YLNNNLLQV	17
469	NLLQVLPPH	17
472	QVLPPHIFS	17
604	AVPLSVLIL	17
610	LILGLLIMF	17
613	GLLIMFITI	17
765	ELQQLGITE	17
768	QLGITEYLR	17
23	PVLSSRGSC	16
163	KVLILNDNA	16
166	ILNDNAIES	16
239	QSIIGDVVC	16
245	VVCNSPPFF	16
284	HLAATSSIN	16
336	KVLSPSGLL	16
420	GNHLTKLSK	16
439	YLYLEYNAI	16
440	LYLEYNAIK	16
502	ILDDLDLLT	16
556	ELKALNSEI	16
559	ALNSEILCP	16
568	GLVNNPSMP	16
597	ILRSLTDAV	16
615	LIMFITIVF	16

Table XXV-V1-HLA-A3-9mers-158P1D7		
621	IVFCAAGIV	16
629	VVLVLHRRR	16
630	VLVLHRRRR	16
650	DNSPVHLQY	16
659	SMYGHKTTH	16
716	HLQRSLLEQ	16
728	SPLTGSNMK	16
769	LGITEYLRK	16
810	KVLVEQTKN	16
812	LVEQTKNEY	16
17	SLHSQTPVL	15
55	MVSEISVPP	15
60	SVPPSRPFQ	15
71	LLNNGLTML	15
110	GLGLLKQLH	15
113	LLKQLHINH	15
116	QLHINHNSL	15
125	EILKEDTFH	15
164	VLILNDNAI	15
232	WLENMPPQS	15
257	ILSRLKKES	15
260	RLKKESICP	15
271	PVYEEHEDP	15
303	ILKLPTKAP	15
369	ILAGNIIHS	15
425	KLSKGMFLG	15
449	EILPGTFNP	15
462	KVLYLNNNL	15
463	VLYLNNNLL	15
473	VLPPHIFSG	15
481	GVPLTKVNL	15
526	VGLQQWIQK	15
626	AGIVVLVLH	15
627	GIVVLVLHR	15
656	LQYSMYGHK	15
707	NEKEGSDAK	15
746	FLSFQDASS	15
788	AHYPGAHEE	15
798	KLMETLMYS	15

Table XXV-V3-HLA-A3-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	SLYEQHMGA	17
7	HMGAHEELK	12

Table XXV-V4-HLA-A3-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
9	ILWSKASGR	23
8	SILWSKASG	16
4	SLMKSILWS	15
5	LMKSILWSK	13
1	IIHSLMKSI	12

Table XXVI-V1-HLA-A26-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
130	DTFHGLENL	32
244	DVVCNSPPF	31
205	EHIGRILDL	27
682	MVSPMVHVY	25
819	EYFELKANL	25
400	EVLEEGSFM	24
498	PVSNILDDL	24
604	AVPLSVLIL	23
761	EKERELQQL	23
148	ITVIEPSAF	22

Table XXVI-V1-HLA-A26-9mers-158P1D7

196	QTLPYVGFL	22
595	DTILRSLTD	22
653	PVHLQYSMY	22
275	EHEDPSGSL	21
453	GTFNPMPKL	21
650	DNSPVHLQY	21
277	EDPSGSLHL	20
336	KVLSPSGLL	20
443	EYNAIKEIL	20
486	KVNLKTNQF	20
520	DCSCDLVGL	20
631	LVLHRRRRY	20
795	EELKLMETL	20
812	LVEQTKNEY	20
87	LTNAISIHL	19
154	SAFSKLNRL	19
182	FVPLTHLDL	19
350	RNIESLSDL	19
462	KVLYLNNNL	19
607	LSVLILGLL	19
610	LILGLLIMF	19
139	EFLQADNNF	18
245	VVCNSPPFF	18
423	LTKLSKGMF	18
481	GVPLTKVNL	18
539	TVTDDILCT	18
628	IVVLVLHRR	18
669	TTERPSASL	18
713	DAKHLQRSL	18
801	ETLMYSRPR	18
106	GAFNGLGLL	17
136	ENLEFLQAD	17
149	TVIEPSAFS	17
225	DLLQLKTWL	17
308	TKAPGLIPY	17
405	GSFMNLTRL	17
410	LTRLQKLYL	17
501	NILDDLDLL	17
590	TTNTADTIL	17
738	KTTNQSTEF	17
739	TTNQSTEFL	17
76	LTMLHTNDF	16
89	NAISIHLGF	16
180	FRFVPLTHL	16

Table XXVI-V1-HLA-A26-9mers-158P1D7

253	FKGSILSRL	16
265	SICPTPPVY	16
298	TKTTSILKL	16
299	KTTSILKLP	16
429	GMFLGLHNL	16
540	VTDDILCTS	16
563	EILCPGLVN	16
593	TADTILRSL	16
815	QTKNEYFEL	16
822	ELKANLHAE	16
58	EISVPPSRP	15
104	EIGAFNGLG	15
133	HGLENLEFL	15
174	SLPPNIFRF	15
250	PPFFKGSIL	15
353	ESLSDLRPP	15
370	LAGNIIHSL	15
378	LMKSDLVEY	15
385	EYFTLEMLH	15
449	EILPGTFNP	15
504	DDLDLLTQI	15
615	LIMFITIVF	15
621	IVFCAAGIV	15
705	ERNEKEGSD	15
725	ENHSPLTGS	15
758	NILEKEREL	15
832	DYLEVLEQQ	15

Table XXVI-V3-HLA-A26-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
5	EQHMGAHEE	10
8	MGAHEELKL	10
6	QHMGAHEEL	8

Table XXVI-V4-HLA-A26-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	IHSLMKSIL	9
5	LMKSILWSK	8
1	IIHSLMKSI	7
4	SLMKSILWS	6
8	SILWSKASG	6
7	KSILWSKAS	5

Table XXVII-V1-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
456	NPMPKLKVL	23
458	MPKLKVLYL	23
692	SPSFGPKHL	22
250	PPFFKGSIL	21
61	VPPSRPFQL	20
278	DPSGSLHLA	20
360	PPPQNPRKL	20
361	PPQNPRKLI	20
517	NPWDCSCDL	20
310	APGLIPYIT	19
175	LPPNIFRFV	18
314	IPYITKPST	18
586	TPATTTNTA	18
306	LPTKAPGLI	17
329	CPIPCNCKV	17
474	LPPHIFSGV	17
625	AAGIVVLVL	17
804	MYSRPRKVL	17
62	PPSRPFQLS	16
237	PPQSIIGDV	16
249	SPPFFKGSI	16
364	NPRKLILAG	16
572	NPSMPTQTS	16
575	MPTQTSYLM	16
652	SPVHLQYSM	16

Table XXVII-V1-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
807	RPRKVLVEQ	16
63	PSRPFQLSL	15
105	IGAFNGLGL	15
159	LNRLKVLIL	15
205	EHIGRILDL	15
207	IGRILDLQL	15
267	CPTPPVYEE	15
316	YITKPSTQL	15
426	LSKGMFLGL	15
602	TDAVPLSVL	15
604	AVPLSVLIL	15
623	FCAAGIVVL	15
752	ASSLYRNIL	15
26	SSRGSCDSL	14
103	IEIGAFNGL	14
152	EPSAFSKLN	14
177	PNIFRFVPL	14
180	FRFVPLTHL	14
221	ACNCDLLQL	14
275	EHEDPSGSL	14
319	KPSTQLPGP	14
326	GPYCPIPCN	14
336	KVLSPSGLL	14
339	SPSGLLIHC	14
410	LTRLQKLYL	14
453	GTFNPMPKL	14
476	PHIFSGVPL	14
520	DCSCDLVGL	14
599	RSLTDAVPL	14
606	PLSVLILGL	14
669	TTERPSASL	14
672	RPSASLYEQ	14
774	YLRKNIAQL	14
830	EPDYLEVLE	14
17	SLHSQTPVL	13
37	CEEKDGTML	13
65	RPFQLSLLN	13
196	QTLPYVGFL	13
198	LPYVGFLEH	13
264	ESICPTPPV	13
277	EDPSGSLHL	13
324	LPGPYCPIP	13
331	IPCNCKVLS	13
359	RPPPQNPRK	13

Table XXVII-V1-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
362	PQNPRKLIL	13
375	IHSLMKSDL	13
402	LEEGSFMNL	13
648	MRDNSPVHL	13
714	AKHLQRSLL	13
767	QQLGITEYL	13
791	PGAHEELKL	13
829	AEPDYLEVL	13
59	ISVPPSRPF	12
68	QLSLLNNGL	12
83	DFSGLTNAI	12
109	NGLGLLKQL	12
151	IEPSAFSKL	12
172	IESLPPNIF	12
176	PPNIFRFVP	12
182	FVPLTHLDL	12
187	HLDLRGNQL	12
189	DLRGNQLQT	12
219	KWACNCDLL	12
234	ENMPPQSII	12
296	MSTKTTSIL	12
298	TKTTSILKL	12
305	KLPTKAPGL	12
323	QLPGPYCPI	12
337	VLSPSGLLI	12
386	YFTLEMLHL	12
415	KLYLNGNHL	12
418	LNGNHLTKL	12
424	TKLSKGMFL	12
434	LHNLEYLYL	12
443	EYNAIKEIL	12
451	LPGTFNPMP	12
481	GVPLTKVNL	12
489	LKTNQFTHL	12
497	LPVSNILDD	12
498	PVSNILDDL	12
500	SNILDDLDL	12
552	LDKKELKAL	12
566	CPGLVNNPS	12
624	CAAGIVVLV	12
684	SPMVHVYRS	12
709	KEGSDAKHL	12
739	TTNQSTEFL	12
789	HYPGAHEEL	12

Table XXVII-V1-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
790	YPGAHEELK	12
795	EELKLMETL	12
819	EYFELKANL	12
5	IHLFYSSLL	11
71	LLNNGLTML	11
79	LHTNDFSGL	11
87	LTNAISIHL	11
119	INHNSLEIL	11
127	LKEDTFHGL	11
133	HGLENLEFL	11
157	SKLNRLKVL	11
167	LNDNAIESL	11
190	LRGNQLQTL	11
195	LQTLPYVGF	11
203	FLEHIGRIL	11
218	NKWACNCDL	11
225	DLLQLKTWL	11
253	FKGSILSRL	11
295	RMSTKTTSI	11
300	TTSILKLPT	11
330	PIPCNCKVL	11
350	RNIESLSDL	11
370	LAGNIIHSL	11
394	LGNNRIEVL	11
405	GSFMNLTRL	11
450	ILPGTFNPM	11
455	FNPMPKLKV	11
462	KVLYLNNNL	11
466	LNNNLLQVL	11
475	PPHIFSGVP	11
479	FSGVPLTKV	11
482	VPLTKVNLK	11
495	THLPVSNIL	11
537	KNTVTDDIL	11
544	ILCTSPGHL	11
548	SPGHLDKKE	11
557	LKALNSEIL	11
561	NSEILCPGL	11
574	SMPTQTSYL	11
590	TTNTADTIL	11
593	TADTILRSL	11
597	ILRSLTDAV	11
681	HMVSPMVHV	11
722	LEQENHSPL	11

Table XXVII-V1-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
741	NQSTEFLSF	11
761	EKERELQQL	11
780	AQLQPDMEA	11
783	QPDMEAHYP	11
805	YSRPRKVLV	11
826	NLHAEPDYL	11

Table XXVII-V3-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
6	QHMGAHEEL	13
8	MGAHEELKL	13
2	SLYEQHMGA	6

Table XXVII-V4-HLA-B0702-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	IHSLMKSIL	13
6	MKSILWSKA	8
1	IIHSLMKSI	7
13	KASGRGRRE	6

Table XXVIII-V1-HLA-B08-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
458	MPKLKVLYL	38
159	LNRLKVLIL	28
456	NPMPKLKVL	27
758	NILEKEREL	27
154	SAFSKLNRL	26
187	HLDLRGNQL	26
250	PPFFKGSIL	26
305	KLPTKAPGL	26
556	ELKALNSEI	26
61	VPPSRPFQL	25
713	DAKHLQRSL	24
258	LSRLKKESI	23
774	YLRKNIAQL	23
552	LDKKELKAL	22
157	SKLNRLKVL	21
205	EHIGRILDL	21
638	RYKKKQVDE	21
734	NMKYKTTNQ	21
815	QTKNEYFEL	21
303	ILKLPTKAP	20
424	TKLSKGMFL	20
426	LSKGMFLGL	20
760	LEKERELQQ	20
126	ILKEDTFHG	19
177	PNIFRFVPL	19
394	LGNNRIEVL	19
463	VLYLNNNLL	19
692	SPSFGPKHL	19
796	ELKLMETLM	19
822	ELKANLHAE	19
17	SLHSQTPVL	18
26	SSRGSCDSL	18
38	EEKDGTMLI	18
68	QLSLLNNGL	18
161	RLKVLILND	18
362	PQNPRKLIL	18
408	MNLTRLQKL	18
482	VPLTKVNLK	18
527	GLQQWIQKL	18
606	PLSVLILGL	18
636	RRRYKKKQV	18
696	GPKHLEEEE	18
813	VEQTKNEYF	18

Table XXVIII-V3-HLA-B08-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Table XXVIII-V3-HLA-B08-9mers-158P1D7

Pos	123456789	score
6	QHMGAHEEL	11
2	SLYEQHMGA	10
8	MGAHEELKL	10

Table XXVIII-V4-HLA-B08-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
9	ILWSKASGR	17
1	IIHSLMKSI	12
2	IHSLMKSIL	12
13	KASGRGRRE	12
3	HSLMKSILW	11
5	LMKSILWSK	10
11	WSKASGRGR	10
4	SLMKSILWS	9

Table XXIX-V1-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
275	EHEDPSGSL	24
375	IHSLMKSDL	24
205	EHIGRILDL	23
495	THLPVSNIL	23
5	IHLFYSSLL	22
476	PHIFSGVPL	22
79	LHTNDFSGL	20
434	LHNLEYLYL	20
132	FHGLENLEF	17
623	FCAAGIVVL	17
687	VHVYRSPSF	17
602	TDAVPLSVL	16
18	LHSQTPVLS	15
360	PPPQNPRKL	15
804	MYSRPRKVL	15
105	IGAFNGLGL	14
345	IHCQERNIE	14
392	LHLGNNRIE	14

Table XXIX-V1-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
405	GSFMNLTRL	14
453	GTFNPMPKL	14
456	NPMPKLKVL	14
481	GVPLTKVNL	14
680	QHMVSPMVH	14
758	NILEKEREL	14
774	YLRKNIAQL	14
788	AHYPGAHEE	14
795	EELKLMETL	14
17	SLHSQTPVL	13
59	ISVPPSRPF	13
93	IHLGFNNIA	13
103	IEIGAFNGL	13
186	THLDLRGNQ	13
196	QTLPYVGFL	13
203	FLEHIGRIL	13
316	YITKPSTQL	13
330	PIPCNCKVL	13
347	CQERNIESL	13
362	PQNPRKLIL	13
394	LGNNRIEVL	13
520	DCSCDLVGL	13
527	GLQQWIQKL	13
544	ILCTSPGHL	13
550	GHLDKKELK	13
593	TADTILRSL	13
606	PLSVLILGL	13
625	AAGIVVLVL	13
648	MRDNSPVHL	13
666	THHTTERPS	13
669	TTERPSASL	13
692	SPSFGPKHL	13
726	NHSPLTGSN	13
793	AHEELKLME	13
819	EYFELKANL	13
827	LHAEPDYLE	13
829	AEPDYLEVL	13
37	CEEKDGTML	12
63	PSRPFQLSL	12
106	GAFNGLGLL	12
119	INHNSLEIL	12
127	LKEDTFHGL	12
133	HGLENLEFL	12
151	IEPSAFSKL	12

Table XXIX-V1-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
154	SAFSKLNRL	12
157	SKLNRLKVL	12
174	SLPPNIFRF	12
177	PNIFRFVPL	12
180	FRFVPLTHL	12
207	IGRILDLQL	12
219	KWACNCDLL	12
225	DLLQLKTWL	12
253	FKGSILSRL	12
277	EDPSGSLHL	12
298	TKTTSILKL	12
381	SDLVEYFTL	12
386	YFTLEMLHL	12
402	LEEGSFMNL	12
429	GMFLGLHNL	12
443	EYNAIKEIL	12
466	LNNNLLQVL	12
549	PGHLDKKEL	12
552	LDKKELKAL	12
557	LKALNSEIL	12
561	NSEILCPGL	12
599	RSLTDAVPL	12
662	GHKTTHHTT	12
667	HHTTERPSA	12
698	KHLEEEEER	12
713	DAKHLQRSL	12
722	LEQENHSPL	12
739	TTNQSTEFL	12
752	ASSLYRNIL	12
761	EKERELQQL	12
789	HYPGAHEEL	12
26	SSRGSCDSL	11
61	VPPSRPFQL	11
68	QLSLLNNGL	11
71	LLNNGLTML	11
109	NGLGLLKQL	11
116	QLHINHNSL	11
130	DTFHGLENL	11
159	LNRLKVLIL	11
167	LNDNAIESL	11
172	IESLPPNIF	11
190	LRGNQLQTL	11
288	TSSINDSRM	11
296	MSTKTTSIL	11

Table XXIX-V1-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
305	KLPTKAPGL	11
335	CKVLSPSGL	11
336	KVLSPSGLL	11
350	RNIESLSDL	11
370	LAGNIIHSL	11
410	LTRLQKLYL	11
415	KLYLNGNHL	11
424	TKLSKGMFL	11
426	LSKGMFLGL	11
432	LGLHNLEYL	11
447	IKEILPGTF	11
458	MPKLKVLYL	11
463	VLYLNNNLL	11
498	PVSNILDDL	11
501	NILDDLDLL	11
517	NPWDCSCDL	11
537	KNTVTDDIL	11
590	TTNTADTIL	11
604	AVPLSVLIL	11
633	LHRRRRYKK	11
654	VHLQYSMYG	11
714	AKHLQRSLL	11
715	KHLQRSLLE	11
747	LSFQDASSL	11
767	QQLGITEYL	11
791	PGAHEELKL	11
815	QTKNEYFEL	11
826	NLHAEPDYL	11

Table XXIX-V3-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
6	QHMGAHEEL	22
8	MGAHEELKL	11

Table XXIX-V4-HLA-B1510-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	IHSLMKSIL	24

Table XXX-V1-HLA-B2705-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
180	FRFVPLTHL	27
358	LRPPPQNPR	25
64	SRPFQLSLL	22
190	LRGNQLQTL	22
429	GMFLGLHNL	22
634	HRRRRYKKK	22
648	MRDNSPVHL	22
690	YRSPSFGPK	22
756	YRNILEKER	22
405	GSFMNLTRL	21
637	RRYKKKQVD	21
255	GSILSRLKK	20
350	RNIESLSDL	20
453	GTFNPMPKL	20
527	GLQQWIQKL	20
719	RSLLEQENH	20
763	ERELQQLGI	20
106	GAFNGLGLL	19
359	RPPPQNPRK	19
462	KVLYLNNNL	19
819	EYFELKANL	19
130	DTFHGLENL	18
139	EFLQADNNF	18
154	SAFSKLNRL	18
205	EHIGRILDL	18
225	DLLQLKTWL	18
252	FFKGSILSR	18
481	GVPLTKVNL	18
599	RSLTDAVPL	18
747	LSFQDASSL	18
809	RKVLVEQTK	18
109	NGLGLLKQL	17

Table XXX-V1-HLA-B2705-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
160	NRLKVLILN	17
202	GFLEHIGRI	17
208	GRILDLQLE	17
211	LDLQLEDNK	17
298	TKTTSILKL	17
301	TSILKLPTK	17
316	YITKPSTQL	17
372	GNIIHSLMK	17
411	TRLQKLYLN	17
420	GNHLTKLSK	17
550	GHLDKKELK	17
610	LILGLLIMF	17
623	FCAAGIVVL	17
627	GIVVLVLHR	17
628	IVVLVLHRR	17
635	RRRRYKKKQ	17
636	RRRYKKKQV	17
698	KHLEEEEER	17
754	SLYRNILEK	17
766	LQQLGITEY	17
774	YLRKNIAQL	17
103	IEIGAFNGL	16
125	EILKEDTFH	16
173	ESLPPNIFR	16
174	SLPPNIFRF	16
201	VGFLEHIGR	16
259	SRLKKESIC	16
336	KVLSPSGLL	16
342	GLLIHCQER	16
366	RKLILAGNI	16
390	EMLHLGNNR	16
397	NRIEVLEEG	16
402	LEEGSFMNL	16
415	KLYLNGNHL	16
478	IFSGVPLTK	16
486	KVNLKTNQF	16
495	THLPVSNIL	16
506	LDLLTQIDL	16
526	VGLQQWIQK	16
659	SMYGHKTTH	16
711	GSDAKHLQR	16
728	SPLTGSNMK	16
738	KTTNQSTEF	16
769	LGITEYLRK	16

Table XXX-V1-HLA-B2705-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
795	EELKLMETL	16
5	IHLFYSSLL	15
10	SSLLACISL	15
20	SQTPVLSSR	15
51	KGIKMVSEI	15
57	SEISVPPSR	15
59	ISVPPSRPF	15
63	PSRPFQLSL	15
71	LLNNGLTML	15
86	GLTNAISIH	15
100	IADIEIGAF	15
107	AFNGLGLLK	15
124	LEILKEDTF	15
132	FHGLENLEF	15
153	PSAFSKLNR	15
155	AFSKLNRLK	15
207	IGRILDLQL	15
250	PPFFKGSIL	15
253	FKGSILSRL	15
305	KLPTKAPGL	15
309	KAPGLIPYI	15
311	PGLIPYITK	15
370	LAGNIIHSL	15
399	IEVLEEGSF	15
408	MNLTRLQKL	15
418	LNGNHLTKL	15
440	LYLEYNAIK	15
463	VLYLNNNLL	15
469	NLLQVLPPH	15
482	VPLTKVNLK	15
500	SNILDDLDL	15
547	TSPGHLDKK	15
604	AVPLSVLIL	15
606	PLSVLILGL	15
609	VLILGLLIM	15
625	AAGIVVLVL	15
629	VVLVLHRRR	15
640	KKKQVDEQM	15
664	KTTHHTTER	15
691	RSPSFGPKH	15
708	EKEGSDAKH	15
729	PLTGSNMKY	15
758	NILEKEREL	15
767	QQLGITEYL	15

Table XXX-V1-HLA-B2705-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
11	SLLACISLH	14
26	SSRGSCDSL	14
37	CEEKDGTML	14
68	QLSLLNNGL	14
89	NAISIHLGF	14
110	GLGLLKQLH	14
113	LLKQLHINH	14
133	HGLENLEFL	14
148	ITVIEPSAF	14
150	VIEPSAFSK	14
151	IEPSAFSKL	14
157	SKLNRLKVL	14
159	LNRLKVLIL	14
167	LNDNAIESL	14
172	IESLPPNIF	14
196	QTLPYVGFL	14
198	LPYVGFLEH	14
221	ACNCDLLQL	14
254	KGSILSRLK	14
277	EDPSGSLHL	14
287	ATSSINDSR	14
294	SRMSTKTTS	14
295	RMSTKTTSI	14
335	CKVLSPSGL	14
347	CQERNIESL	14
349	ERNIESLSD	14
360	PPPQNPRKL	14
365	PRKLILAGN	14
368	LILAGNIIH	14
375	IHSLMKSDL	14
381	SDLVEYFTL	14
394	LGNNRIEVL	14
414	QKLYLNGNH	14
417	YLNGNHLTK	14
424	TKLSKGMFL	14
456	NPMPKLKVL	14
458	MPKLKVLYL	14
476	PHIFSGVPL	14
546	CTSPGHLDK	14
552	LDKKELKAL	14
573	PSMPTQTSY	14
598	LRSLTDAVP	14
602	TDAVPLSVL	14
607	LSVLILGLL	14

Table XXX-V1-HLA-B2705-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
626	AGIVVLVLH	14
630	VLVLHRRRR	14
652	SPVHLQYSM	14
669	TTERPSASL	14
687	VHVYRSPSF	14
701	EEEEERNEK	14
707	NEKEGSDAK	14
713	DAKHLQRSL	14
778	NIAQLQPDM	14
791	PGAHEELKL	14
792	GAHEELKLM	14
802	TLMYSRPRK	14
806	SRPRKVLVE	14
4	WIHLFYSSL	13
32	DSLCNCEEK	13
46	INCEAKGIK	13
87	LTNAISIHL	13
95	LGFNNIADI	13
111	LGLLKQLHI	13
119	INHNSLEIL	13
143	ADNNFITVI	13
177	PNIFRFVPL	13
183	VPLTHLDLR	13
187	HLDLRGNQL	13
192	GNQLQTLPY	13
195	LQTLPYVGF	13
244	DVVCNSPPF	13
275	EHEDPSGSL	13
291	INDSRMSTK	13
296	MSTKTTSIL	13
308	TKAPGLIPY	13
312	GLIPYITKP	13
362	PQNPRKLIL	13
384	VEYFTLEML	13
385	EYFTLEMLH	13
386	YFTLEMLHL	13
391	MLHLGNNRI	13
400	EVLEEGSFM	13
404	EGSFMNLTR	13
407	FMNLTRLQK	13
410	LTRLQKLYL	13
423	LTKLSKGMF	13
426	LSKGMFLGL	13
432	LGLHNLEYL	13

Table XXX-V1-HLA-B2705-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
433	GLHNLEYLY	13
434	LHNLEYLYL	13
447	IKEILPGTF	13
457	PMPKLKVLY	13
466	LNNNLLQVL	13
471	LQVLPPHIF	13
501	NILDDLDLL	13
504	DDLDLLTQI	13
529	QQWIQKLSK	13
537	KNTVTDDIL	13
549	PGHLDKKEL	13
567	PGLVNNPSM	13
590	TTNTADTIL	13
593	TADTILRSL	13
611	ILGLLIMFI	13
613	GLLIMFITI	13
615	LIMFITIVF	13
633	LHRRRRYKK	13
705	ERNEKEGSD	13
709	KEGSDAKHL	13
714	AKHLQRSLL	13
718	QRSLLEQEN	13
739	TTNQSTEFL	13
741	NQSTEFLSF	13
749	FQDASSLYR	13
752	ASSLYRNIL	13
761	EKERELQQL	13
789	HYPGAHEEL	13
799	LMETLMYSR	13
801	ETLMYSRPR	13
808	PRKVLVEQT	13
812	LVEQTKNEY	13
829	AEPDYLEVL	13

Table XXX-V3-HLA-B2705-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
8	MGAHEELKL	14
6	QHMGAHEEL	13
3	LYEQHMGAH	10
7	HMGAHEELK	10
1	ASLYEQHMG	6

Table XXX-V4-HLA-B2705-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	IHSLMKSIL	14
5	LMKSILWSK	14
9	ILWSKASGR	14
12	SKASGRGRR	14
11	WSKASGRGR	11
1	IIHSLMKSI	9
4	SLMKSILWS	7
7	KSILWSKAS	6
8	SILWSKASG	6
13	KASGRGRRE	6

Table XXXI-V1-HLA-B2709-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
636	RRRYKKKQV	23
180	FRFVPLTHL	22
648	MRDNSPVHL	21
64	SRPFQLSLL	20
190	LRGNQLQTL	20
599	RSLTDAVPL	19
763	ERELQQLGI	19
366	RKLILAGNI	16
405	GSFMNLTRL	16
429	GMFLGLHNL	16
453	GTFNPMPKL	16
637	RRYKKKQVD	16
106	GAFNGLGLL	15
208	GRILDLQLE	15
336	KVLSPSGLL	15
350	RNIESLSDL	15

Table XXXI-V1-HLA-B2709-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
462	KVLYLNNNL	15
481	GVPLTKVNL	15
709	KEGSDAKHL	15
154	SAFSKLNRL	14
196	QTLPYVGFL	14
202	GFLEHIGRI	14
221	ACNCDLLQL	14
305	KLPTKAPGL	14
415	KLYLNGNHL	14
635	RRRRYKKKQ	14
747	LSFQDASSL	14
5	IHLFYSSLL	13
109	NGLGLLKQL	13
130	DTFHGLENL	13
207	IGRILDLQL	13
253	FKGSILSRL	13
411	TRLQKLYLN	13
424	TKLSKGMFL	13
495	THLPVSNIL	13
500	SNILDDLDL	13
501	NILDDLDLL	13
527	GLQQWIQKL	13
537	KNTVTDDIL	13
604	AVPLSVLIL	13
613	GLLIMFITI	13
625	AAGIVVLVL	13
767	QQLGITEYL	13
819	EYFELKANL	13
10	SSLLACISL	12
17	SLHSQTPVL	12
51	KGIKMVSEI	12
61	VPPSRPFQL	12
63	PSRPFQLSL	12
79	LHTNDFSGL	12
89	NAISIHLGF	12
103	IEIGAFNGL	12
105	IGAFNGLGL	12
133	HGLENLEFL	12
151	IEPSAFSKL	12
157	SKLNRLKVL	12
159	LNRLKVLIL	12
160	NRLKVLILN	12
171	AIESLPPNI	12
177	PNIFRFVPL	12
205	EHIGRILDL	12

Table XXXI-V1-HLA-B2709-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
219	KWACNCDLL	12
225	DLLQLKTWL	12
250	PPFFKGSIL	12
259	SRLKKESIC	12
277	EDPSGSLHL	12
295	RMSTKTTSI	12
298	TKTTSILKL	12
316	YITKPSTQL	12
362	PQNPRKLIL	12
381	SDLVEYFTL	12
384	VEYFTLEML	12
386	YFTLEMLHL	12
408	MNLTRLQKL	12
432	LGLHNLEYL	12
458	MPKLKVLYL	12
463	VLYLNNNLL	12
476	PHIFSGVPL	12
506	LDLLTQIDL	12
520	DCSCDLVGL	12
607	LSVLILGLL	12
621	IVFCAAGIV	12
671	ERPSASLYE	12
758	NILEKEREL	12
775	LRKNIAQLQ	12
795	EELKLMETL	12
806	SRPRKVLVE	12
808	PRKVLVEQT	12
16	ISLHSQTPV	11
27	SRGSCDSLC	11
37	CEEKDGTML	11
59	ISVPPSRPF	11
70	SLLNNGLTM	11
85	SGLTNAISI	11
87	LTNAISIHL	11
111	LGLLKQLHI	11
119	INHNSLEIL	11
139	EFLQADNNF	11
158	KLNRLKVLI	11
182	FVPLTHLDL	11
187	HLDLRGNQL	11
193	NQLQTLPYV	11
203	FLEHIGRIL	11
294	SRMSTKTTS	11
296	MSTKTTSIL	11
309	KAPGLIPYI	11

Table XXXI-V1-HLA-B2709-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
335	CKVLSPSGL	11
349	ERNIESLSD	11
358	LRPPPQNPR	11
365	PRKLILAGN	11
367	KLILAGNII	11
370	LAGNIIHSL	11
375	IHSLMKSDL	11
397	NRIEVLEEG	11
402	LEEGSFMNL	11
410	LTRLQKLYL	11
426	LSKGMFLGL	11
434	LHNLEYLYL	11
443	EYNAIKEIL	11
456	NPMPKLKVL	11
486	KVNLKTNQF	11
489	LKTNQFTHL	11
498	PVSNILDDL	11
504	DDLDLLTQI	11
544	ILCTSPGHL	11
549	PGHLDKKEL	11
561	NSEILCPGL	11
567	PGLVNNPSM	11
593	TADTILRSL	11
603	DAVPLSVLI	11
606	PLSVLILGL	11
608	SVLILGLLI	11
623	FCAAGIVVL	11
624	CAAGIVVLV	11
640	KKKQVDEQM	11
675	ASLYEQHMV	11
681	HMVSPMVHV	11
690	YRSPSFGPK	11
714	AKHLQRSLL	11
738	KTTNQSTEF	11
752	ASSLYRNIL	11
761	EKERELQQL	11
774	YLRKNIAQL	11
791	PGAHEELKL	11
792	GAHEELKLM	11
803	LMYSRPRKV	11
828	HAEPDYLEV	11
829	AEPDYLEVL	11

Table XXXI-V3-HLA-B2709-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
8	MGAHEELKL	11
6	QHMGAHEEL	10

Table XXXI-V4-HLA-B2709-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
2	IHSLMKSIL	11
1	IIHSLMKSI	10

Table XXXII-V1-HLA-B4402-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
829	AEPDYLEVL	27
103	IEIGAFNGL	26
124	LEILKEDTF	25
670	TERPSASLY	25
172	IESLPPNIF	24
442	LEYNAIKEI	24
709	KEGSDAKHL	24
795	EELKLMETL	24
38	EEKDGTMLI	23
151	IEPSAFSKL	23
402	LEEGSFMNL	22
205	EHIGRILDL	21
384	VEYFTLEML	21
399	IEVLEEGSF	21
722	LEQENHSPL	21
813	VEQTKNEYF	21
37	CEEKDGTML	20
174	SLPPNIFRF	19
233	LENMPPQSI	19
456	NPMPKLKVL	19

Table XXXII-V1-HLA-B4402-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
157	SKLNRLKVL	18
109	NGLGLLKQL	17
562	SEILCPGLV	17
604	AVPLSVLIL	17
682	MVSPMVHVY	17
752	ASSLYRNIL	17
89	NAISIHLGF	16
100	IADIEIGAF	16
143	ADNNFITVI	16
164	VLILNDNAI	16
177	PNIFRFVPL	16
221	ACNCDLLQL	16
224	CDLLQLKTW	16
265	SICPTPPVY	16
298	TKTTSILKL	16
370	LAGNIIHSL	16
394	LGNNRIEVL	16
500	SNILDDLDL	16
625	AAGIVVLVL	16
650	DNSPVHLQY	16
703	EEERNEKEG	16
714	AKHLQRSLL	16
804	MYSRPRKVL	16
818	NEYFELKAN	16
48	CEAKGIKMV	15
57	SEISVPPSR	15
95	LGFNNIADI	15
106	GAFNGLGLL	15
154	SAFSKLNRL	15
167	LNDNAIESL	15
187	HLDLRGNQL	15
196	QTLPYVGFL	15
276	HEDPSGSLH	15
308	TKAPGLIPY	15
330	PIPCNCKVL	15
347	CQERNIESL	15
360	PPPQNPRKL	15
362	PQNPRKLIL	15
408	MNLTRLQKL	15
409	NLTRLQKLY	15
429	GMFLGLHNL	15
448	KEILPGTFN	15
486	KVNLKTNQF	15
495	THLPVSNIL	15
501	NILDDLDLL	15

Table XXXII-V1-HLA-B4402-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
552	LDKKELKAL	15
593	TADTILRSL	15
606	PLSVLILGL	15
615	LIMFITIVF	15
623	FCAAGIVVL	15
692	SPSFGPKHL	15
741	NQSTEFLSF	15
761	EKERELQQL	15
774	YLRKNIAQL	15
10	SSLLACISL	14
59	ISVPPSRPF	14
61	VPPSRPFQL	14
63	PSRPFQLSL	14
64	SRPFQLSLL	14
76	LTMLHTNDF	14
83	DFSGLTNAI	14
85	SGLTNAISI	14
128	KEDTFHGLE	14
135	LENLEFLQA	14
138	LEFLQADNN	14
139	EFLQADNNF	14
234	ENMPPQSII	14
277	EDPSGSLHL	14
305	KLPTKAPGL	14
309	KAPGLIPYI	14
337	VLSPSGLLI	14
350	RNIESLSDL	14
367	KLILAGNII	14
403	EEGSFMNLT	14
405	GSFMNLTRL	14
415	KLYLNGNHL	14
453	GTFNPMPKL	14
463	VLYLNNNLL	14
476	PHIFSGVPL	14
498	PVSNILDDL	14
527	GLQQWIQKL	14
555	KELKALNSE	14
573	PSMPTQTSY	14
574	SMPTQTSYL	14
599	RSLTDAVPL	14
607	LSVLILGLL	14
610	LILGLLIMF	14
631	LVLHRRRRY	14
648	MRDNSPVHL	14
701	EEEEERNEK	14

Table XXXII-V1-HLA-B4402-9mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.		
Pos	123456789	score
702	EEEERNEKE	14
744	TEFLSFQDA	14
766	LQQLGITEY	14
767	QQLGITEYL	14
819	EYFELKANL	14
825	ANLHAEPDY	14
1	MKLWIHLFY	13
17	SLHSQTPVL	13
51	KGIKMVSEI	13
68	QLSLLNNGL	13
127	LKEDTFHGL	13
130	DTFHGLENL	13
133	HGLENLEFL	13
148	ITVIEPSAF	13
159	LNRLKVLIL	13
180	FRFVPLTHL	13
182	FVPLTHLDL	13
190	LRGNQLQTL	13
192	GNQLQTLPY	13
204	LEHIGRILD	13
212	DLQLEDNKW	13
219	KWACNCDLL	13
263	KESICPTPP	13
274	EEHEDPSGS	13
275	EHEDPSGSL	13
336	KVLSPSGLL	13
348	QERNIESLS	13
352	IESLSDLRP	13
361	PPQNPRKLI	13
379	MKSDLVEYF	13
381	SDLVEYFTL	13
389	LEMLHLGNN	13
418	LNGNHLTKL	13
426	LSKGMFLGL	13
432	LGLHNLEYL	13
443	EYNAIKEIL	13
457	PMPKLKVLY	13
458	MPKLKVLYL	13
462	KVLYLNNNL	13
466	LNNNLLQVL	13
471	LQVLPPHIF	13
481	GVPLTKVNL	13
506	LDLLTQIDL	13
511	QIDLEDNPW	13
520	DCSCDLVGL	13

Table XXXII-V1-HLA-B4402-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
523	CDLVGLQQW	13
549	PGHLDKKEL	13
603	DAVPLSVLI	13
704	EERNEKEGS	13
707	NEKEGSDAK	13
724	QENHSPLTG	13
747	LSFQDASSL	13
748	SFQDASSLY	13
758	NILEKEREL	13
760	LEKERELQQ	13
772	TEYLRKNIA	13
786	MEAHYPGAH	13
797	LKLMETLMY	13

Table XXXII-V3-HLA-B4402-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
6	QHMGAHEEL	12
8	MGAHEELKL	12
4	YEQHMGAHE	10
1	ASLYEQHMG	5

Table XXXII-V4-HLA-B4402-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
3	HSLMKSILW	13
2	IHSLMKSIL	12
1	IIHSLMKSI	10
7	KSILWSKAS	9
4	SLMKSILWS	6
14	ASGRGRREE	6

Table XXXIII-V1-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
603	DAVPLSVLI	25
751	DASSLYRNI	25
306	LPTKAPGLI	24
625	AAGIVVLVL	24
111	LGLLKQLHI	23
175	LPPNIFRFV	23
309	KAPGLIPYI	23
456	NPMPKLKVL	23
142	QADNNFITV	22
474	LPPHIFSGV	22
624	CAAGIVVLV	22
85	SGLTNAISI	21
154	SAFSKLNRL	21
249	SPPFFKGSI	21
329	CPIPCNCKV	21
360	PPPQNPRKL	21
361	PPQNPRKLI	21
458	MPKLKVLYL	21
713	DAKHLQRSL	21
51	KGIKMVSEI	20
95	LGFNNIADI	20
593	TADTILRSL	20
61	VPPSRPFQL	19
237	PPQSIIGDV	19
370	LAGNIIHSL	19
504	DDLDLLTQI	19
517	NPWDCSCDL	19
692	SPSFGPKHL	19
828	HAEPDYLEV	19
106	GAFNGLGLL	18
109	NGLGLLKQL	18
198	LPYVGFLEH	18
250	PPFFKGSIL	18
394	LGNNRIEVL	18
442	LEYNAIKEI	18
482	VPLTKVNLK	18
803	LMYSRPRKV	18
133	HGLENLEFL	17
278	DPSGSLHLA	17
314	IPYITKPST	17
432	LGLHNLEYL	17
439	YLYLEYNAI	17
605	VPLSVLILG	17
613	GLLIMFITI	17
83	DFSGLTNAI	16

Table XXXIII-V1-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
202	GFLEHIGRI	16
586	TPATTTNTA	16
105	IGAFNGLGL	15
143	ADNNFITVI	15
170	NAIESLPPN	15
183	VPLTHLDLR	15
207	IGRILDLQL	15
236	MPPQSIIGD	15
283	LHLAATSSI	15
285	LAATSSIND	15
326	GPYCPIPCN	15
524	DLVGLQQWI	15
589	TTTNTADTI	15
601	LTDAVPLSV	15
791	PGAHEELKL	15
807	RPRKVLVEQ	15
13	LACISLHSQ	14
16	ISLHSQTPV	14
45	LINCEAKGI	14
49	EAKGIKMVS	14
74	NGLTMLHTN	14
140	FLQADNNFI	14
269	TPPVYEEHE	14
339	SPSGLLIHC	14
364	NPRKLILAG	14
391	MLHLGNNRI	14
445	NAIKEILPG	14
451	LPGTFNPMP	14
470	LLQVLPPHI	14
497	LPVSNILDD	14
532	IQKLSKNTV	14
558	KALNSEILC	14
566	CPGLVNNPS	14
587	PATTTNTAD	14
622	VFCAAGIVV	14
728	SPLTGSNMK	14
792	GAHEELKLM	14
22	TPVLSSRGS	13
100	IADIEIGAF	13
157	SKLNRLKVL	13
176	PPNIFRFVP	13
193	NQLQTLPYV	13
199	PYVGFLEHI	13
225	DLLQLKTWL	13
233	LENMPPQSI	13

Table XXXIII-V1-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
258	LSRLKKESI	13
286	AATSSINDS	13
295	RMSTKTTSI	13
298	TKTTSILKL	13
324	LPGPYCPIP	13
331	IPCNCKVLS	13
337	VLSPSGLLI	13
344	LIHCQERNI	13
359	RPPPQNPRK	13
366	RKLILAGNI	13
384	VEYFTLEML	13
408	MNLTRLQKL	13
455	FNPMPKLKV	13
463	VLYLNNNLL	13
475	PPHIFSGVP	13
479	FSGVPLTKV	13
494	FTHLPVSNI	13
536	SKNTVTDDI	13
548	SPGHLDKKE	13
549	PGHLDKKEL	13
572	NPSMPTQTS	13
608	SVLILGLLI	13
611	ILGLLIMFI	13
623	FCAAGIVVL	13
672	RPSASLYEQ	13
684	SPMVHVYRS	13
758	NILEKEREL	13
771	ITEYLRKNI	13
779	IAQLQPDME	13
790	YPGAHEELK	13
829	AEPDYLEVL	13
8	FYSSLLACI	12
41	DGTMLINCE	12
53	IKMVSEISV	12
65	RPFQLSLLN	12
89	NAISIHLGF	12
92	SIHLGFNNI	12
97	FNNIADIEI	12
130	DTFHGLENL	12
151	IEPSAFSKL	12
152	EPSAFSKLN	12
156	FSKLNRLKV	12
159	LNRLKVLIL	12
164	VLILNDNAI	12
267	CPTPPVYEE	12

Table XXXIII-V1-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
319	KPSTQLPGP	12
323	QLPGPYCPI	12
415	KLYLNGNHL	12
418	LNGNHLTKL	12
426	LSKGMFLGL	12
465	YLNNNLLQV	12
466	LNNNLLQVL	12
495	THLPVSNIL	12
506	LDLLTQIDL	12
520	DCSCDLVGL	12
544	ILCTSPGHL	12
556	ELKALNSEI	12
575	MPTQTSYLM	12
614	LLIMFITIV	12
620	TIVFCAAGI	12
621	IVFCAAGIV	12
674	SASLYEQHM	12

Table XXXIII-V3-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
8	MGAHEELKL	16
6	QHMGAHEEL	7

Table XXXIII-V4-HLA-B5101-9mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
1	IIHSLMKSI	13
13	KASGRGRRE	13
2	IHSLMKSIL	9

Table XXXIV-V1-HLA-A1-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
669	TTERPSASLY	33
307	PTKAPGLIPY	25
430	MFLGLHNLEY	23
796	ELKLMETLMY	23
191	RGNQLQTLPY	21
435	HNLEYLYLEY	21
456	NPMPKLKVLY	21
649	RDNSPVHLQY	21
743	STEFLSFQDA	21
747	LSFQDASSLY	21
134	GLENLEFLQA	20
150	VIEPSAFSKL	20
264	ESICPTPPVY	20
276	HEDPSGSLHL	20
728	SPLTGSNMKY	20
781	QLQPDMEAHY	20
203	FLEHIGRILD	19
820	YFELKANLHA	19
377	SLMKSDLVEY	18
630	VLVLHRRRRY	18
652	SPVHLQYSMY	18
805	YSRPRKVLVE	18
128	KEDTFHGLEN	17
408	MNLTRLQKLY	17
432	LGLHNLEYLY	17
502	ILDDLDLLTQ	17
518	PWDCSCDLVG	17
540	VTDDILCTSP	17
601	LTDAVPLSVL	17
681	HMVSPMVHVY	17
759	ILEKERELQQ	17
811	VLVEQTKNEY	17
830	EPDYLEVLEQ	17
297	STKTTSILKL	16
317	ITKPSTQLPG	16
351	NIESLSDLRP	16
561	NSEILCPGLV	16
723	EQENHSPLTG	16
765	ELQQLGITEY	16
771	ITEYLRKNIA	16

Table XXXIV-V3-HLA-A1-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
4	LYEQHMGAHE	11
8	HMGAHEELKL	8
2	ASLYEQHMGA	5

Table XXXIV-V4-HLA-A1-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
4	HSLMKSILWS	10
3	IHSLMKSILW	6
12	WSKASGRGRR	5
8	KSILWSKASG	4

Table XXXV-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
369	ILAGNIIHSL	33
417	YLNGNHLTKL	31
166	ILNDNAIESL	30
70	SLLNNGLTML	28
158	KLNRLKVLIL	27
189	DLRGNQLQTL	27
465	YLNNNLLQVL	27
613	GLLIMFITIV	27
407	FMNLTRLQKL	26
610	LILGLLIMFI	26
126	ILKEDTFHGL	25
431	FLGLHNLEYL	25
600	SLTDAVPLSV	25
174	SLPPNIFRFV	24
393	HLGNNRIEVL	24

Table XXXV-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
473	VLPPHIFSGV	24
551	HLDKKELKAL	24
94	HLGFNNIADI	23
118	HINHNSLEIL	23
425	KLSKGMFLGL	23
441	YLEYNAIKEI	23
592	NTADTILRSL	23
624	CAAGIVVLVL	23
150	VIEPSAFSKL	22
257	ILSRLKKESI	22
282	SLHLAATSSI	22
297	STKTTSILKL	22
343	LLIHCQERNI	22
401	VLEEGSFMNL	22
433	GLHNLEYLYL	22
746	FLSFQDASSL	22
802	TLMYSRPRKV	22
12	LLACISLHSQ	21
78	MLHTNDFSGL	21
377	SLMKSDLVEY	21
469	NLLQVLPPHI	21
531	WIQKLSKNTV	21
581	YLMVTTPATT	21
596	TILRSLTDAV	21
606	PLSVLILGLL	21
647	QMRDNSPVHL	21
721	LLEQENHSPL	21
44	MLINCEAKGI	20
52	GIKMVSEISV	20
86	GLTNAISIHL	20
110	GLGLLKQLHI	20
374	IIHSLMKSDL	20
409	NLTRLQKLYL	20
457	PMPKLKVLYL	20
478	IFSGVPLTKV	20
502	ILDDLDLLTQ	20
601	LTDAVPLSVL	20
603	DAVPLSVLIL	20
803	LMYSRPRKVL	20
206	HIGRILDLQL	19
220	WACNCDLLQL	19
232	WLENMPPQSI	19
305	KLPTKAPGLI	19
464	LYLNNNLLQV	19

Table XXXV-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
488	NLKTNQFTHL	19
505	DLDLLTQIDL	19
526	VGLQQWIQKL	19
543	DILCTSPGHL	19
564	ILCPGLVNNP	19
605	VPLSVLILGL	19
616	IMFITIVFCA	19
619	ITIVFCAAGI	19
623	FCAAGIVVLV	19
668	HTTERPSASL	19
676	SLYEQHMVSP	19
720	SLLEQENHSP	19
754	SLYRNILEKE	19
827	LHAEPDYLEV	19
828	HAEPDYLEVL	19
4	WIHLFYSSLL	18
15	CISLHSQTPV	18
60	SVPPSRPFQL	18
102	DIEIGAFNGL	18
240	SIIGDVVCNS	18
295	RMSTKTTSIL	18
304	LKLPTKAPGL	18
337	VLSPSGLLIH	18
346	HCQERNIESL	18
382	DLVEYFTLEM	18
383	LVEYFTLEML	18
392	LHLGNNRIEV	18
500	SNILDDLDLL	18
7	LFYSSLLACI	17
104	EIGAFNGLGL	17
105	IGAFNGLGLL	17
141	LQADNNFITV	17
163	KVLILNDNAI	17
170	NAIESLPPNI	17
204	LEHIGRILDL	17
260	RLKKESICPT	17
308	TKAPGLIPYI	17
415	KLYLNGNHLT	17
462	KVLYLNNNLL	17
490	KTNQFTHLPV	17
519	WDCSCDLVGL	17
559	ALNSEILCPG	17
608	SVLILGLLIM	17
609	VLILGLLIMF	17

Table XXXV-V1-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
620	TIVFCAAGIV	17
621	IVFCAAGIVV	17
622	VFCAAGIVVL	17
674	SASLYEQHMV	17
760	LEKERELQQL	17
770	GITEYLRKNI	17
788	AHYPGAHEEL	17
798	KLMETLMYSR	17
3	LWIHLFYSSL	16
6	HLFYSSLLAC	16
50	AKGIKMVSEI	16
99	NIADIEIGAF	16
113	LLKQLHINHN	16
115	KQLHINHNSL	16
142	QADNNFITVI	16
192	GNQLQTLPYV	16
252	FFKGSILSRL	16
313	LIPYITKPST	16
336	KVLSPSGLLI	16
368	LILAGNIIHS	16
390	EMLHLGNNRI	16
412	RLQKLYLNGN	16
428	KGMFLGLHNL	16
445	NAIKEILPGT	16
454	TFNPMPKLKV	16
480	SGVPLTKVNL	16
494	FTHLPVSNIL	16
497	LPVSNILDDL	16
501	NILDDLDLLT	16
556	ELKALNSEIL	16
560	LNSEILCPGL	16
582	LMVTTPATTT	16
611	ILGLLIMFIT	16
615	LIMFITIVFC	16
712	SDAKHLQRSL	16
811	VLVEQTKNEY	16

Table XXXV-V3-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
8	HMGAHEELKL	21
3	SLYEQHMGAH	16

Table XXXV-V4-HLA-A2-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
2	IIHSLMKSIL	20
1	NIIHSLMKSI	19
5	SLMKSILWSK	19
6	LMKSILWSKA	15
9	SILWSKASGR	13
10	ILWSKASGRG	13
14	KASGRGRREE	9

Table XXXVI-V1-HLA-A0203-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
278	DPSGSLHLAA	19
617	MFITIVFCAA	19
279	PSGSLHLAAT	17
618	FITIVFCAAG	17
5	IHLFYSSLLA	10
41	DGTMLINCEA	10
81	TNDFSGLTNA	10
92	SIHLGFNNIA	10
98	NNIADIEIGA	10
134	GLENLEFLQA	10
146	NFITVIEPSA	10
162	LKVLILNDNA	10
212	DLQLEDNKWA	10
277	EDPSGSLHLA	10
301	TSILKLPTKA	10
362	PQNPRKLILA	10

Table XXXVI-V1-HLA-A0203-10mers-158P1D7

437	LEYLYLEYNA	10
550	GHLDKKELKA	10
579	TSYLMVTTPA	10
585	TTPATTTNTA	10
595	DTILRSLTDA	10
616	IMFITIVFCA	10
666	THHTTERPSA	10
705	ERNEKEGSDA	10
743	STEFLSFQDA	10
771	ITEYLRKNIA	10
779	IAQLQPDMEA	10
784	PDMEAHYPGA	10
816	TKNEYFELKA	10
820	YFELKANLHA	10
6	HLFYSSLLAC	9
42	GTMLINCEAK	9
82	NDFSGLTNAI	9
93	IHLGFNNIAD	9
99	NIADIEIGAF	9
135	LENLEFLQAD	9
147	FITVIEPSAF	9
163	KVLILNDNAI	9
213	LQLEDNKWAC	9
302	SILKLPTKAP	9
363	QNPRKLILAG	9
438	EYLYLEYNAI	9
551	HLDKKELKAL	9
580	SYLMVTTPAT	9
586	TPATTTNTAD	9
596	TILRSLTDAV	9
667	HHTTERPSAS	9
706	RNEKEGSDAK	9
744	TEFLSFQDAS	9
772	TEYLRKNIAQ	9
780	AQLQPDMEAH	9
785	DMEAHYPGAH	9
817	KNEYFELKAN	9
821	FELKANLHAE	9

Table XXXVI-V3-HLA-A0203-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
2	ASLYEQHMGA	10
3	SLYEQHMGAH	9
4	LYEQHMGAHE	8

Table XXXVI-V4-HLA-A0203-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
6	LMKSILWSKA	10
7	MKSILWSKAS	9
8	KSILWSKASG	8

Table XXXVII-V1-HLA-A3-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
149	TVIEPSAFSK	29
439	YLYLEYNAIK	28
290	SINDSRMSTK	27
477	HIFSGVPLTK	26
768	QLGITEYLRK	26
525	LVGLQQWIQK	24
632	VLHRRRRYKK	24
781	QLQPDMEAHY	24
178	NIFRFVPLTH	23
210	ILDLQLEDNK	23
446	AIKEILPGTF	23
631	LVLHRRRRYK	23
245	VVCNSPPFFK	22
597	ILRSLTDAVP	22
676	SLYEQHMVSP	22
729	PLTGSNMKYK	22
796	ELKLMETLMY	22
336	KVLSPSGLLI	21
367	KLILAGNIIH	21
377	SLMKSDLVEY	21
481	GVPLTKVNLK	21
614	LLIMFITIVF	21
655	HLQYSMYGHK	21
682	MVSPMVHVYR	21
123	SLEILKEDTF	20
194	QLQTLPYVGF	20
337	VLSPSGLLIH	20
357	DLRPPPQNPR	20
416	LYLNGNHLTK	20
502	ILDDLDLLTQ	20
798	KLMETLMYSR	20
45	LINCEAKGIK	19

Table XXXVII-V1-HLA-A3-10mers-158P1D7		
158	KLNRLKVLIL	19
189	DLRGNQLQTL	19
398	RIEVLEEGSF	19
406	SFMNLTRLQK	19
472	QVLPPHIFSG	19
609	VLILGLLIMF	19
621	IVFCAAGIVV	19
11	SLLACISLHS	18
23	PVLSSRGSCD	18
60	SVPPSRPFQL	18
254	KGSILSRLKK	18
310	APGLIPYITK	18
371	AGNIIHSLMK	18
415	KLYLNGNHLT	18
463	VLYLNNNLLQ	18
581	YLMVTTPATT	18
600	SLTDAVPLSV	18
630	VLVLHRRRRY	18
746	FLSFQDASSL	18
754	SLYRNILEKE	18
759	ILEKERELQQ	18
44	MLINCEAKGI	17
106	GAFNGLGLLK	17
134	GLENLEFLQA	17
147	FITVIEPSAF	17
163	KVLILNDNAI	17
164	VLILNDNAIE	17
197	TLPYVGFLEH	17
206	HIGRILDLQL	17
257	ILSRLKKESI	17
265	SICPTPPVYE	17
282	SLHLAATSSI	17
303	ILKLPTKAPG	17
369	ILAGNIIHSL	17
608	SVLILGLLIM	17
628	IVVLVLHRRR	17
629	VVLVLHRRRR	17
688	HVYRSPSFGP	17
765	ELQQLGITEY	17
811	VLVEQTKNEY	17
2	KLWIHLFYSS	16
17	SLHSQTPVLS	16
70	SLLNNGLTML	16
71	LLNNGLTMLH	16
99	NIADIEIGAF	16
104	EIGAFNGLGL	16
112	GLLKQLHINH	16
116	QLHINHNSLE	16
171	AIESLPPNIF	16
214	QLEDNKWACN	16
312	GLIPYITKPS	16
409	NLTRLQKLYL	16

Table XXXVII-V1-HLA-A3-10mers-158P1D7		
422	HLTKLSKGMF	16
425	KLSKGMFLGL	16
473	VLPPHIFSGV	16
633	LHRRRRYKKK	16
649	RDNSPVHLQY	16
686	MVHVYRSPSF	16
716	HLQRSLLEQE	16
720	SLLEQENHSP	16
753	SSLYRNILEK	16
774	YLRKNIAQLQ	16
822	ELKANLHAEP	16
90	AISIHLGFNN	15
161	RLKVLILNDN	15
166	ILNDNAIESL	15
182	FVPLTHLDLR	15
209	RILDLQLEDN	15
244	DVVCNSPPFF	15
260	RLKKESICPT	15
271	PVYEEHEDPS	15
300	TTSILKLPTK	15
305	KLPTKAPGLI	15
314	IPYITKPSTQ	15
393	HLGNNRIEVL	15
419	NGNHLTKLSK	15
450	ILPGTFNPMP	15
462	KVLYLNNNLL	15
465	YLNNNLLQVL	15
470	LLQVLPPHIF	15
507	DLLTQIDLED	15
528	LQQWIQKLSK	15
539	TVTDDILCTS	15
544	ILCTSPGHLD	15
562	SEILCPGLVN	15
564	ILCPGLVNNP	15
569	LVNNPSMPTQ	15
583	MVTTPATTTN	15
669	TTERPSASLY	15
706	RNEKEGSDAK	15
808	PRKVLVEQTK	15
810	KVLVEQTKNE	15
6	HLFYSSLLAC	14
68	QLSLLNNGLT	14
75	GLTMLHTNDF	14
110	GLGLLKQLHI	14
126	ILKEDTFHGL	14
150	VIEPSAFSKL	14
165	LILNDNAIES	14
174	SLPPNIFRFV	14
200	YVGFLEHIGR	14
226	LLQLKTWLEN	14
228	QLKTWLENMP	14
232	WLENMPPQSI	14

Table XXXVII-V1-HLA-A3-10mers-158P1D7		
240	SIIGDVVCNS	14
264	ESICPTPPVY	14
323	QLPGPYCPIP	14
382	DLVEYFTLEM	14
400	EVLEEGSFMN	14
412	RLQKLYLNGN	14
433	GLHNLEYLYL	14
460	KLKVLYLNNN	14
483	PLTKVNLKTN	14
486	KVNLKTNQFT	14
501	NILDDLDLLT	14
534	KLSKNTVTDD	14
563	EILCPGLVNN	14
596	TILRSLTDAV	14
604	AVPLSVLILG	14
613	GLLIMFITIV	14
643	QVDEQMRDNS	14
689	VYRSPSFGPK	14
812	LVEQTKNEYF	14
815	QTKNEYFELK	14

Table XXXVII-V3-HLA-A3-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
3	SLYEQHMGAH	22
7	QHMGAHEELK	14

Table XXXVII-V4-HLA-A3-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
5	SLMKSILWSK	23
9	SILWSKASGR	21
1	NIIHSLMKSI	13
2	IIHSLMKSIL	13
10	ILWSKASGRG	13
8	KSILWSKASG	12

Table XXXVIII-V1-HLA-A26-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
244	DVVCNSPPFF	30
603	DAVPLSVLIL	26
104	EIGAFNGLGL	24
264	ESICPTPPVY	24
595	DTILRSLTDA	24
765	ELQQLGITEY	24
129	EDTFHGLENL	23
173	ESLPPNIFRF	23
297	STKTTSILKL	23
307	PTKAPGLIPY	23
349	ERNIESLSDL	23
383	LVEYFTLEML	23
385	EYFTLEMLHL	23
400	EVLEEGSFMN	23
773	EYLRKNIAQL	23
58	EISVPPSRPF	22
274	EEHEDPSGSL	22
404	EGSFMNLTRL	22
592	NTADTILRSL	22
796	ELKLMETLMY	22
60	SVPPSRPFQL	21
189	DLRGNQLQTL	21
543	DILCTSPGHL	21
601	LTDAVPLSVL	21
102	DIEIGAFNGL	20
130	DTFHGLENLE	20
668	HTTERPSASL	20
669	TTERPSASLY	20
99	NIADIEIGAF	19
681	HMVSPMVHVY	19
686	MVHVYRSPSF	19
814	EQTKNEYFEL	19
149	TVIEPSAFSK	18
205	EHIGRILDLQ	18
462	KVLYLNNNLL	18
539	TVTDDILCTS	18
556	ELKALNSEIL	18
563	EILCPGLVNN	18
589	TTTNTADTIL	18
609	VLILGLLIMF	18
708	EKEGSDAKHL	18
801	ETLMYSRPRK	18
812	LVEQTKNEYF	18
423	LTKLSKGMFL	17
497	LPVSNILDDL	17

Table XXXVIII-V1-HLA-A26-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
500	SNILDDLDLL	17
505	DLDLLTQIDL	17
538	NTVTDDILCT	17
652	SPVHLQYSMY	17
713	DAKHLQRSLL	17
738	KTTNQSTEFL	17
751	DASSLYRNIL	17
55	MVSEISVPPS	16
118	HINHNSLEIL	16
217	DNKWACNCDL	16
446	AIKEILPGTF	16
472	QVLPPHIFSG	16
494	FTHLPVSNIL	16
516	DNPWDCSCDL	16
608	SVLILGLLIM	16
621	IVFCAAGIVV	16
811	VLVEQTKNEY	16
819	EYFELKANLH	16
39	EKDGTMLINC	15
147	FITVIEPSAF	15
150	VIEPSAFSKL	15
277	EDPSGSLHLA	15
346	HCQERNIESL	15
377	SLMKSDLVEY	15
382	DLVEYFTLEM	15
449	EILPGTFNPM	15
604	AVPLSVLILG	15
671	ERPSASLYEQ	15
747	LSFQDASSLY	15
760	LEKERELQQL	15
763	ERELQQLGIT	15
830	EPDYLEVLEQ	15
3	LWIHLFYSSL	14
63	PSRPFQLSLL	14
70	SLLNNGLTML	14
125	EILKEDTFHG	14
166	ILNDNAIESL	14
181	RFVPLTHLDL	14
182	FVPLTHLDLR	14
195	LQTLPYVGFL	14
206	HIGRILDLQL	14
220	WACNCDLLQL	14
300	TTSILKLPTK	14
374	IIHSLMKSDL	14
398	RIEVLEEGSF	14

Table XXXVIII-V1-HLA-A26-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
480	SGVPLTKVNL	14
481	GVPLTKVNLK	14
485	TKVNLKTNQF	14
546	CTSPGHLDKK	14
605	VPLSVLILGL	14
628	IVVLVLHRRR	14
630	VLVLHRRRRY	14
705	ERNEKEGSDA	14

Table XXXVIII-V3-HLA-A26-10mers-158P1D

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
6	EQHMGAHEEL	18
8	HMGAHEELKL	10

Table XXXVIII-V4-HLA-A26-10mers-158P1D

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
2	IIHSLMKSIL	14
1	NIIHSLMKSI	9
5	SLMKSILWSK	6
9	SILWSKASGR	6

Table XXXIX-V1-HLA-B0702-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
62	PPSRPFQLSL	24
176	PPNIFRFVPL	24

Table XXXIX-V1-HLA-B0702-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
790	YPGAHEELKL	24
278	DPSGSLHLAA	23
475	PPHIFSGVPL	23
329	CPIPCNCKVL	22
359	RPPPQNPRKL	22
361	PPQNPRKLIL	22
605	VPLSVLILGL	22
548	SPGHLDKKEL	21
807	RPRKVLVEQT	21
249	SPPFFKGSIL	20
497	LPVSNILDDL	20
482	VPLTKVNLKT	18
566	CPGLVNNPSM	18
575	MPTQTSYLMV	18
237	PPQSIIGDVV	17
360	PPPQNPRKLI	17
425	KLSKGMFLGL	17
624	CAAGIVVLVL	17
152	EPSAFSKLNR	16
198	LPYVGFLEHI	16
236	MPPQSIIGDV	16
517	NPWDCSCDLV	16
104	EIGAFNGLGL	15
598	LRSLTDAVPL	15
830	EPDYLEVLEQ	15
16	ISLHSQTPVL	14
155	AFSKLNRLKV	14
158	KLNRLKVLIL	14
179	IFRFVPLTHL	14
181	RFVPLTHLDL	14
189	DLRGNQLQTL	14
276	HEDPSGSLHL	14
295	RMSTKTTSIL	14
319	KPSTQLPGPY	14
331	IPCNCKVLSP	14
339	SPSGLLIHCQ	14
364	NPRKLILAGN	14
369	ILAGNIIHSL	14
404	EGSFMNLTRL	14
456	NPMPKLKVLY	14
457	PMPKLKVLYL	14
603	DAVPLSVLIL	14
622	VFCAAGIVVL	14
647	QMRDNSPVHL	14
672	RPSASLYEQH	14

Table XXXIX-V1-HLA-B0702-10mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
63	PSRPFQLSLL	13
65	RPFQLSLLNN	13
206	HIGRILDLQL	13
306	LPTKAPGLIP	13
310	APGLIPYITK	13
324	LPGPYCPIPC	13
385	EYFTLEMLHL	13
417	YLNGNHLTKL	13
480	SGVPLTKVNL	13
551	HLDKKELKAL	13
560	LNSEILCPGL	13
572	NPSMPTQTSY	13
573	PSMPTQTSYL	13
586	TPATTTNTAD	13
601	LTDAVPLSVL	13
708	EKEGSDAKHL	13
738	KTTNQSTEFL	13
751	DASSLYRNIL	13
788	AHYPGAHEEL	13
9	YSSLLACISL	12
25	LSSRGSCDSL	12
105	IGAFNGLGLL	12
126	ILKEDTFHGL	12
132	FHGLENLEFL	12
150	VIEPSAFSKL	12
175	LPPNIFRFVP	12
183	VPLTHLDLRG	12
195	LQTLPYVGFL	12
204	LEHIGRILDL	12
220	WACNCDLLQL	12
252	FFKGSILSRL	12
263	KESICPTPPV	12
297	STKTTSILKL	12
304	LKLPTKAPGL	12
380	KSDLVEYFTL	12
393	HLGNNRIEVL	12
409	NLTRLQKLYL	12
428	KGMFLGLHNL	12
433	GLHNLEYLYL	12
451	LPGTFNPMPK	12
478	IFSGVPLTKV	12
488	NLKTNQFTHL	12
499	VSNILDDLDL	12
519	WDCSCDLVGL	12
556	ELKALNSEIL	12

Table XXXIX-V1-HLA-B0702-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
606	PLSVLILGLL	12
692	SPSFGPKHLE	12
712	SDAKHLQRSL	12
746	FLSFQDASSL	12
773	EYLRKNIAQL	12
783	QPDMEAHYPG	12
803	LMYSRPRKVL	12
814	EQTKNEYFEL	12
825	ANLHAEPDYL	12
828	HAEPDYLEVL	12
22	TPVLSSRGSC	11
36	NCEEKDGTML	11
60	SVPPSRPFQL	11
61	VPPSRPFQLS	11
70	SLLNNGLTML	11
78	MLHTNDFSGL	11
102	DIEIGAFNGL	11
108	FNGLGLLKQL	11
115	KQLHINHNSL	11
129	EDTFHGLENL	11
153	PSAFSKLNRL	11
156	FSKLNRLKVL	11
166	ILNDNAIESL	11
218	NKWACNCDLL	11
224	CDLLQLKTWL	11
267	CPTPPVYEEH	11
274	EEHEDPSGSL	11
314	IPYITKPSTQ	11
315	PYITKPSTQL	11
349	ERNIESLSDL	11
374	IIHSLMKSDL	11
401	VLEEGSFMNL	11
423	LTKLSKGMFL	11
431	FLGLHNLEYL	11
452	PGTFNPMPKL	11
455	FNPMPKLKVL	11
462	KVLYLNNNLL	11
465	YLNNNLLQVL	11
474	LPPHIFSGVP	11
505	DLDLLTQIDL	11
589	TTTNTADTIL	11
592	NTADTILRSL	11
623	FCAAGIVVLV	11
668	HTTERPSASL	11
684	SPMVHVYRSP	11

Table XXXIX-V1-HLA-B0702-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
691	RSPSFGPKHL	11
713	DAKHLQRSLL	11
721	LLEQENHSPL	11
757	RNILEKEREL	11
762	KERELQQLGI	11
766	LQQLGITEYL	11
818	NEYFELKANL	11

Table XXXIX-V3-HLA-B0702-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
8	HMGAHEELKL	14
6	EQHMGAHEEL	11
2	ASLYEQHMGA	8
9	MGAHEELKLM	7

Table XXXIX-V4-HLA-B0702-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
2	IIHSLMKSIL	11
1	NIIHSLMKSI	6
6	LMKSILWSKA	6
14	KASGRGRREE	6

Table XL-V1-HLA-B08-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XL-V3-HLA-B08-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XL-V4-HLA-B08-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLI-V1-HLA-B1510-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLI-V3-HLA-B1510-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLI-V4-HLA-B1510-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLII-V1-HLA-B2705-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLII-V3-HLA-B2705-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLII-V4-HLA-B2705-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLIII-V1-HLA-B2709-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLIII-V3-HLA-B2709-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLIII-V4-HLA-B2709-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLIV-V1-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
276	HEDPSGSLHL	25
138	LEFLQADNNF	24
204	LEHIGRILDL	24
274	EEHEDPSGSL	22
760	LEKERELQQL	22
794	HEELKLMETL	22
442	LEYNAIKEIL	21
818	NEYFELKANL	21
37	CEEKDGTMLI	20
555	KELKALNSEI	20
762	KERELQQLGI	20
173	ESLPPNIFRF	19
329	CPIPCNCKVL	19
233	LENMPPQSII	18
773	EYLRKNIAQL	18
60	SVPPSRPFQL	17
99	NIADIEIGAF	17
223	NCDLLQLKTW	17
264	ESICPTPPVY	17
297	STKTTSILKL	17
359	RPPPQNPRKL	17
456	NPMPKLKVLY	17
500	SNILDDLDLL	17
562	SEILCPGLVN	17
829	AEPDYLEVLE	17
385	EYFTLEMLHL	16
448	KEILPGTFNP	16
551	HLDKKELKAL	16
609	VLILGLLIMF	16
614	LLIMFITIVF	16
708	EKEGSDAKHL	16
788	AHYPGAHEEL	16
44	MLINCEAKGI	15
57	SEISVPPSRP	15
63	PSRPFQLSLL	15
82	NDFSGLTNAI	15
103	IEIGAFNGLG	15
108	FNGLGLLKQL	15
150	VIEPSAFSKL	15
156	FSKLNRLKVL	15
171	AIESLPPNIF	15
304	LKLPTKAPGL	15
315	PYITKPSTQL	15
369	ILAGNIIHSL	15

Table XLIV-V1-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
393	HLGNNRIEVL	15
408	MNLTRLQKLY	15
446	AIKEILPGTF	15
455	FNPMPKLKVL	15
480	SGVPLTKVNL	15
510	TQIDLEDNPW	15
522	SCDLVGLQQW	15
526	VGLQQWIQKL	15
573	PSMPTQTSYL	15
603	DAVPLSVLIL	15
605	VPLSVLILGL	15
622	VFCAAGIVVL	15
744	TEFLSFQDAS	15
757	RNILEKEREL	15
765	ELQQLGITEY	15
796	ELKLMETLMY	15
821	FELKANLHAE	15
825	ANLHAEPDYL	15
828	HAEPDYLEVL	15
38	EEKDGTMLIN	14
58	EISVPPSRPF	14
70	SLLNNGLTML	14
124	LEILKEDTFH	14
128	KEDTFHGLEN	14
135	LENLEFLQAD	14
142	QADNNFITVI	14
151	IEPSAFSKLN	14
158	KLNRLKVLIL	14
166	ILNDNAIESL	14
181	RFVPLTHLDL	14
186	THLDLRGNQL	14
201	VGFLEHIGRI	14
220	WACNCDLLQL	14
308	TKAPGLIPYI	14
319	KPSTQLPGPY	14
352	IESLSDLRPP	14
377	SLMKSDLVEY	14
380	KSDLVEYFTL	14
403	EEGSFMNLTR	14
404	EGSFMNLTRL	14
425	KLSKGMFLGL	14
428	KGMFLGLHNL	14
438	EYLYLEYNAI	14
462	KVLYLNNNLL	14

Table XLIV-V1-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
485	TKVNLKTNQF	14
588	ATTTNTADTI	14
592	NTADTILRSL	14
598	LRSLTDAVPL	14
624	CAAGIVVLVL	14
670	TERPSASLYE	14
691	RSPSFGPKHL	14
702	EEEERNEKEG	14
728	SPLTGSNMKY	14
772	TEYLRKNIAQ	14
795	EELKLMETLM	14
803	LMYSRPRKVL	14
3	LWIHLFYSSL	13
9	YSSLLACISL	13
16	ISLHSQTPVL	13
62	PPSRPFQLSL	13
91	ISIHLGFNNI	13
104	EIGAFNGLGL	13
115	KQLHINHNSL	13
117	LHINHNSLEI	13
129	EDTFHGLENL	13
147	FITVIEPSAF	13
157	SKLNRLKVLI	13
163	KVLILNDNAI	13
170	NAIESLPPNI	13
172	IESLPPNIFR	13
189	DLRGNQLQTL	13
206	HIGRILDLQL	13
211	LDLQLEDNKW	13
215	LEDNKWACNC	13
248	NSPPFFKGSI	13
263	KESICPTPPV	13
295	RMSTKTTSIL	13
305	KLPTKAPGLI	13
346	HCQERNIESL	13
349	ERNIESLSDL	13
360	PPPQNPRKLI	13
389	LEMLHLGNNR	13
402	LEEGSFMNLT	13
407	FMNLTRLQKL	13
409	NLTRLQKLYL	13
417	YLNGNHLTKL	13
430	MFLGLHNLEY	13
441	YLEYNAIKEI	13

Table XLIV-V1-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
457	PMPKLKVLYL	13
465	YLNNNLLQVL	13
488	NLKTNQFTHL	13
505	DLDLLTQIDL	13
548	SPGHLDKKEL	13
601	LTDAVPLSVL	13
606	PLSVLILGLL	13
610	LILGLLIMFI	13
612	LGLLIMFITI	13
630	VLVLHRRRRY	13
647	QMRDNSPVHL	13
669	TTERPSASLY	13
681	HMVSPMVHVY	13
701	EEEEERNEKE	13
703	EEERNEKEGS	13
704	EERNEKEGSD	13
709	KEGSDAKHLQ	13
738	KTTNQSTEFL	13
747	LSFQDASSLY	13
751	DASSLYRNIL	13
764	RELQQLGITE	13
781	QLQPDMEAHY	13
4	WIHLFYSSLL	12
25	LSSRGSCDSL	12
50	AKGIKMVSEI	12
67	FQLSLLNNGL	12
75	GLTMLHTNDF	12
78	MLHTNDFSGL	12
86	GLTNAISIHL	12
102	DIEIGAFNGL	12
105	IGAFNGLGLL	12
123	SLEILKEDTF	12
126	ILKEDTFHGL	12
131	TFHGLENLEF	12
132	FHGLENLEFL	12
139	EFLQADNNFI	12
153	PSAFSKLNRL	12
176	PPNIFRFVPL	12
191	RGNQLQTLPY	12
194	QLQTLPYVGF	12
195	LQTLPYVGFL	12
202	GFLEHIGRIL	12
218	NKWACNCDLL	12
224	CDLLQLKTWL	12

Table XLIV-V1-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
249	SPPFFKGSIL	12
252	FFKGSILSRL	12
307	PTKAPGLIPY	12
322	TQLPGPYCPI	12
334	NCKVLSPSGL	12
335	CKVLSPSGLL	12
336	KVLSPSGLLI	12
343	LLIHCQERNI	12
348	QERNIESLSD	12
361	PPQNPRKLIL	12
390	EMLHLGNNRI	12
414	QKLYLNGNHL	12
431	FLGLHNLEYL	12
432	LGLHNLEYLY	12
433	GLHNLEYLYL	12
435	HNLEYLYLEY	12
461	LKVLYLNNNL	12
470	LLQVLPPHIF	12
494	FTHLPVSNIL	12
503	LDDLDLLTQI	12
514	LEDNPWDCSC	12
519	WDCSCDLVGL	12
536	SKNTVTDDIL	12
543	DILCTSPGHL	12
556	ELKALNSEIL	12
572	NPSMPTQTSY	12
619	ITIVFCAAGI	12
649	RDNSPVHLQY	12
700	LEEEEERNEK	12
707	NEKEGSDAKH	12
712	SDAKHLQRSL	12
713	DAKHLQRSLL	12
740	TNQSTEFLSF	12
746	FLSFQDASSL	12
766	LQQLGITEYL	12
790	YPGAHEELKL	12
800	METLMYSRPR	12
814	EQTKNEYFEL	12
824	KANLHAEPDY	12

Table XLIV-V3-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
5	YEQHMGAHEE	12
6	EQHMGAHEEL	12
8	HMGAHEELKL	12

Table XLIV-V4-HLA-B4402-10mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.		
Pos	1234567890	score
1	NIIHSLMKSI	14
3	IHSLMKSILW	14
2	IIHSLMKSIL	10

Table XLV-V1-HLA-B5101-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLV-V3-HLA-B5101-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLV-V4-HLA-B5101-10mers-158P1D7		
Pos	1234567890	score
NoResultsFound.		

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
6	HLFYSSLLACISLHS	34
300	TTSILKLPTKAPGLI	33
73	NNGLTMLHTNDFSGL	32

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
554	KKELKALNSEILCPG	31
744	TEFLSFQDASSLYRN	31
145	NNFITVIEPSAFSKL	30
169	DNAIESLPPNIFRFV	30
468	NNLLQVLPPHIFSGV	30
153	PSAFSKLNRLKVLIL	29
444	YNAIKEILPGTFNPM	29
15	CISLHSQTPVLSSRG	28
42	GTMLINCEAKGIKMV	28
177	PNIFRFVPLTHLDLR	28
230	KTWLENMPPQSIIGD	28
467	NNNLLQVLPPHIFSG	28
572	NPSMPTQTSYLMVTT	28
606	PLSVLILGLLIMFIT	28
121	HNSLEILKEDTFHGL	27
129	EDTFHGLENLEFLQA	27
161	RLKVLILNDNAIESL	27
179	IFRFVPLTHLDLRGN	27
200	YVGFLEHIGRILDLQ	27
364	NPRKLILAGNIIHSL	27
383	LVEYFTLEMLHLGNN	27
420	GNHLTKLSKGMFLGL	27
436	NLEYLYLEYNAIKEI	27
491	TNQFTHLPVSNILDD	27
1	MKLWIHLFYSSLLAC	26
81	TNDFSGLTNAISIHL	26
102	DIEIGAFNGLGLLKQ	26
192	GNQLQTLPYVGFLEH	26
452	PGTFNPMPKLKVLYL	26
455	FNPMPKLKVLYLNNN	26
476	PHIFSGVPLTKVNLK	26
529	QQWIQKLSKNTVTDD	26
595	DTILRSLTDAVPLSV	26
611	ILGLLIMFITIVFCA	26
618	FITIVFCAAGIVVLV	26
817	KNEYFELKANLHAEP	26
19	HSQTPVLSSRGSCDS	25
94	HLGFNNIADIEIGAF	25
108	FNGLGLLKQLHINHN	25
132	FHGLENLEFLQADNN	25
135	LENLEFLQADNNFIT	25
156	FSKLNRLKVLILNDN	25
279	PSGSLHLAATSSIND	25
313	LIPYITKPSTQLPGP	25
314	IPYITKPSTQLPGPY	25

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
332	PCNCKVLSPSGLLIH	25
388	TLEMLHLGNNRIEVL	25
396	NNRIEVLEEGSFMNL	25
407	FMNLTRLQKLYLNGN	25
431	FLGLHNLEYLYLEYN	25
441	YLEYNAIKEILPGTF	25
503	LDDLDLLTQIDLEDN	25
551	HLDKKELKALNSEIL	25
559	ALNSEILCPGLVNNP	25
50	AKGIKMVSEISVPPS	24
144	DNNFITVIEPSAFSK	24
163	KVLILNDNAIESLPP	24
184	PLTHLDLRGNQLQTL	24
229	LKTWLENMPPQSIIG	24
255	GSILSRLKKESICPT	24
334	NCKVLSPSGLLIHCQ	24
349	ERNIESLSDLRPPPQ	24
363	QNPRKLILAGNIIHS	24
372	GNIIHSLMKSDLVEY	24
412	RLQKLYLNGNHLTKL	24
460	KLKVLYLNNNLLQVL	24
604	AVPLSVLILGLLIMF	24
605	VPLSVLILGLLIMFI	24
608	SVLILGLLIMFITIV	24
615	LIMFITIVFCAAGIV	24
619	ITIVFCAAGIVVLVL	24
645	DEQMRDNSPVHLQYS	24
686	MVHVYRSPSFGPKHL	24
724	QENHSPLTGSNMKYK	24
797	LKLMETLMYSRPRKV	24
800	METLMYSRPRKVLVE	24
2	KLWIHLFYSSLLACI	23
22	TPVLSSRGSCDSLCN	23
52	GIKMVSEISVPPSRP	23
56	VSEISVPPSRPFQLS	23
84	FSGLTNAISIHLGFN	23
97	FNNIADIEIGAFNGL	23
242	IGDVVCNSPPFFKGS	23
280	SGSLHLAATSSINDS	23
310	APGLIPYITKPSTQL	23
380	KSDLVEYFTLEMLHL	23
483	PLTKVNLKTNQFTHL	23
578	QTSYLMVTTPATTTN	23
598	LRSLTDAVPLSVLIL	23
612	LGLLIMFITIVFCAA	23

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
683	VSPMVHVYRSPSFGP	23
718	QRSLLEQENHSPLTG	23
732	GSNMKYKTTNQSTEF	23
780	AQLQPDMEAHYPGAH	23
794	HEELKLMETLMYSRP	23
9	YSSLLACISLHSQTP	22
12	LLACISLHSQTPVLS	22
49	EAKGIKMVSEISVPP	22
53	IKMVSEISVPPSRPF	22
58	EISVPPSRPFQLSLL	22
166	ILNDNAIESLPPNIF	22
204	LEHIGRILDLQLEDN	22
223	NCDLLQLKTWLENMP	22
235	NMPPQSIIGDVVCNS	22
239	QSIIGDVVCNSPPFF	22
293	DSRMSTKTTSILKLP	22
303	ILKLPTKAPGLIPYI	22
352	IESLSDLRPPPQNPR	22
357	DLRPPPQNPRKLILA	22
541	TDDILCTSPGHLDKK	22
577	TQTSYLMVTTPATTT	22
594	ADTILRSLTDAVPLS	22
641	KKQVDEQMRDNSPVH	22
674	SASLYEQHMVSPMVH	22
684	SPMVHVYRSPSFGPK	22
776	RKNIAQLQPDMEAHY	22
100	IADIEIGAFNGLGLL	21
105	IGAFNGLGLLKQLHI	21
260	RLKKESICPTPPVYE	21
373	NIIHSLMKSDLVEYF	21
487	VNLKTNQFTHLPVSN	21
651	NSPVHLQYSMYGHKT	21
736	KYKTTNQSTEFLSFQ	21
55	MVSEISVPPSRPFQL	20
182	FVPLTHLDLRGNQLQ	20
198	LPYVGFLEHIGRILD	20
410	LTRLQKLYLNGNHLT	20
423	LTKLSKGMFLGLHNL	20
445	NAIKEILPGTFNPMP	20
472	QVLPPHIFSGVPLTK	20
497	LPVSNILDDLDLLTQ	20
549	PGHLDKKELKALNSE	20
569	LVNNPSMPTQTSYLM	20
676	SLYEQHMVSPMVHVY	20
760	LEKERELQQLGITEY	20

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
772	TEYLRKNIAQLQPDM	20
88	TNAISIHLGFNNIAD	19
124	LEILKEDTFHGLENL	19
250	PPFFKGSILSRLKKE	19
304	LKLPTKAPGLIPYIT	19
397	NRIEVLEEGSFMNLT	19
405	GSFMNLTRLQKLYLN	19
415	KLYLNGNHLTKLSKG	19
438	EYLYLEYNAIKEILP	19
473	VLPPHIFSGVPLTKV	19
614	LLIMFITIVFCAAGI	19
620	TIVFCAAGIVVLVLH	19
753	SSLYRNILEKERELQ	19
793	AHEELKLMETLMYSR	19
818	NEYFELKANLHAEPD	19
5	IHLFYSSLLACISLH	18
13	LACISLHSQTPVLSS	18
39	EKDGTMLINCEAKGI	18
65	RPFQLSLLNNGLTML	18
68	QLSLLNNGLTMLHTN	18
76	LTMLHTNDFSGLTNA	18
137	NLEFLQADNNFITVI	18
146	NFITVIEPSAFSKLN	18
187	HLDLRGNQLQTLPYV	18
210	ILDLQLEDNKWACNC	18
227	LQLKTWLENMPPQSI	18
286	AATSSINDSRMSTKT	18
302	SILKLPTKAPGLIPY	18
404	EGSFMNLTRLQKLYL	18
421	NHLTKLSKGMFLGLH	18
426	LSKGMFLGLHNLEYL	18
428	KGMFLGLHNLEYLYL	18
462	KVLYLNNNLLQVLPP	18
465	YLNNNLLQVLPPHIF	18
471	LQVLPPHIFSGVPLT	18
481	GVPLTKVNLKTNQFT	18
486	KVNLKTNQFTHLPVS	18
580	SYLMVTTPATTTNTA	18
592	NTADTILRSLTDAVP	18
616	IMFITIVFCAAGIVV	18
617	MFITIVFCAAGIVVL	18
675	ASLYEQHMVSPMVHV	18
703	EEERNEKEGSDAKHL	18
743	STEFLSFQDASSLYR	18
763	ERELQQLGITEYLRK	18

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
768	QLGITEYLRKNIAQL	18
771	ITEYLRKNIAQLQPD	18
802	TLMYSRPRKVLVEQT	18
7	LFYSSLLACISLHSQ	17
34	LCNCEEKDGTMLINC	17
35	CNCEEKDGTMLINCE	17
44	MLINCEAKGIKMVSE	17
66	PFQLSLLNNGLTMLH	17
67	FQLSLLNNGLTMLHT	17
82	NDFSGLTNAISIHLG	17
89	NAISIHLGFNNIADI	17
90	AISIHLGFNNIADIE	17
92	SIHLGFNNIADIEIG	17
111	LGLLKQLHINHNSLE	17
116	QLHINHNSLEILKED	17
148	ITVIEPSAFSKLNRL	17
159	LNRLKVLILNDNAIE	17
164	VLILNDNAIESLPPN	17
172	IESLPPNIFRFVPLT	17
226	LLQLKTWLENMPPQS	17
247	CNSPPFFKGSILSRL	17
254	KGSILSRLKKESICP	17
257	ILSRLKKESICPTPP	17
261	LKKESICPTPPVYEE	17
278	DPSGSLHLAATSSIN	17
299	KTTSILKLPTKAPGL	17
318	TKPSTQLPGPYCPIP	17
341	SGLLIHCQERNIESL	17
376	HSLMKSDLVEYFTLE	17
386	YFTLEMLHLGNNRIE	17
419	NGNHLTKLSKGMFLG	17
429	GMFLGLHNLEYLYLE	17
439	YLYLEYNAIKEILPG	17
458	MPKLKVLYLNNNLLQ	17
463	VLYLNNNLLQVLPPH	17
464	LYLNNNLLQVLPPHI	17
478	IFSGVPLTKVNLKTN	17
522	SCDLVGLQQWIQKLS	17
525	LVGLQQWIQKLSKNT	17
528	LQQWIQKLSKNTVTD	17
537	KNTVTDDILCTSPGH	17
539	TVTDDILCTSPGHLD	17
546	CTSPGHLDKKELKAL	17
576	PTQTSYLMVTTPATT	17
586	TPATTTNTADTILRS	17

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
596	TILRSLTDAVPLSVL	17
601	LTDAVPLSVLILGLL	17
607	LSVLILGLLIMFITI	17
609	VLILGLLIMFITIVF	17
625	AAGIVVLVLHRRRRY	17
626	AGIVVLVLHRRRRYK	17
627	GIVVLVLHRRRRYKK	17
637	RRYKKKQVDEQMRDN	17
706	RNEKEGSDAKHLQRS	17
711	GSDAKHLQRSLLEQE	17
741	NQSTEFLSFQDASSL	17
801	ETLMYSRPRKVLVEQ	17
820	YFELKANLHAEPDYL	17
21	QTPVLSSRGSCDSLC	16
110	GLGLLKQLHINHNSL	16
123	SLEILKEDTFHGLEN	16
142	QADNNFITVIEPSAF	16
147	FITVIEPSAFSKLNR	16
160	NRLKVLILNDNAIES	16
207	IGRILDLQLEDNKWA	16
222	CNCDLLQLKTWLENM	16
233	LENMPPQSIIGDVVC	16
238	PQSIIGDVVCNSPPF	16
248	NSPPFFKGSILSRLK	16
269	TPPVYEEHEDPSGSL	16
271	PVYEEHEDPSGSLHL	16
319	KPSTQLPGPYCPIPC	16
321	STQLPGPYCPIPCNC	16
325	PGPYCPIPCNCKVLS	16
328	YCPIPCNCKVLSPSG	16
333	CNCKVLSPSGLLIHC	16
366	RKLILAGNIIHSLMK	16
367	KLILAGNIIHSLMKS	16
369	ILAGNIIHSLMKSDL	16
378	LMKSDLVEYFTLEML	16
381	SDLVEYFTLEMLHLG	16
391	MLHLGNNRIEVLEEG	16
395	GNNRIEVLEEGSFMN	16
399	IEVLEEGSFMNLTRL	16
402	LEEGSFMNLTRLQKL	16
434	LHNLEYLYLEYNAIK	16
446	AIKEILPGTFNPMPK	16
447	IKEILPGTFNPMPKL	16
448	KEILPGTFNPMPKLK	16
500	SNILDDLDLLTQIDL	16

Table XLVI-V1-HLA-DRB-0101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
511	QIDLEDNPWDCSCDL	16
519	WDCSCDLVGLQQWIQ	16
542	DDILCTSPGHLDKKE	16
558	KALNSEILCPGLVNN	16
562	SEILCPGLVNNPSMP	16
563	EILCPGLVNNPSMPT	16
581	YLMVTTPATTTNTAD	16
603	DAVPLSVLILGLLIM	16
658	YSMYGHKTTHHTTER	16
671	ERPSASLYEQHMVSP	16
689	VYRSPSFGPKHLEEE	16
719	RSLLEQENHSPLTGS	16
735	MKYKTTNQSTEFLSF	16
746	FLSFQDASSLYRNIL	16
749	FQDASSLYRNILEKE	16
769	LGITEYLRKNIAQLQ	16
810	KVLVEQTKNEYFELK	16
821	FELKANLHAEPDYLE	16

Table XLVI-V3-HLA-DRB-0101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
7	ASLYEQHMGAHEELK	18
11	EQHMGAHEELKLMET	18
3	ERPSASLYEQHMGAH	16
8	SLYEQHMGAHEELKL	15
6	SASLYEQHMGAHEEL	14
5	PSASLYEQHMGAHEE	10
12	QHMGAHEELKLMETL	10
9	LYEQHMGAHEELKLM	9
14	MGAHEELKLMETLMY	8

Table XLVI-V4-HLA-DRB-0101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
10	SLMKSILWSKASGRG	26
5	GNIIHSLMKSILWSK	24
9	HSLMKSILWSKASGR	24
14	SILWSKASGRGRREE	23
12	MKSILWSKASGRGRR	18
4	AGNIIHSLMKSILWS	17
2	ILAGNIIHSLMKSIL	16
1	LILAGNIIHSLMKSI	14
13	KSILWSKASGRGRRE	14
6	NIIHSLMKSILWSKA	13
3	LAGNIIHSLMKSILW	12

Table XLVII-V1-HLA-DRB-0301-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
779	IAQLQPDMEAHYPGA	36
376	HSLMKSDLVEYFTLE	31
124	LEILKEDTFHGLENL	30
460	KLKVLYLNNNLLQVL	28
809	RKVLVEQTKNEYFEL	27
138	LEFLQADNNFITVIE	26
407	FMNLTRLQKLYLNGN	26
420	GNHLTKLSKGMFLGL	26
628	IVVLVLHRRRRYKKK	26
801	ETLMYSRPRKVLVEQ	26
121	HNSLEILKEDTFHGL	25
372	GNIIHSLMKSDLVEY	25
396	NNRIEVLEEGSFMNL	25
428	KGMFLGLHNLEYLYL	25
499	VSNILDDLDLLTQID	25
503	LDDLDLLTQIDLEDN	25
810	KVLVEQTKNEYFELK	25
129	EDTFHGLENLEFLQA	24
163	KVLILNDNAIESLPP	22
238	PQSIIGDVVCNSPPF	22
794	HEELKLMETLMYSRP	22
68	QLSLLNNGLTMLHTN	21
73	NNGLTMLHTNDFSGL	21
145	NNFITVIEPSAFSKL	21

Table XLVII-V1-HLA-DRB-0301-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
169	DNAIESLPPNIFRFV	21
399	IEVLEEGSFMNLTRL	21
405	GSFMNLTRLQKLYLN	21
444	YNAIKEILPGTFNPM	21
498	PVSNILDDLDLLTQI	21
537	KNTVTDDILCTSPGH	21
541	TDDILCTSPGHLDKK	21
607	LSVLILGLLIMFITI	21
645	DEQMRDNSPVHLQYS	21
756	YRNILEKERELQQLG	21
2	KLWIHLFYSSLLACI	20
41	DGTMLINCEAKGIKM	20
97	FNNIADIEIGAFNGL	20
148	ITVIEPSAFSKLNRL	20
156	FSKLNRLKVLILNDN	20
185	LTHLDLRGNQLQTLP	20
187	HLDLRGNQLQTLPYV	20
192	GNQLQTLPYVGFLEH	20
204	LEHIGRILDLQLEDN	20
206	HIGRILDLQLEDNKW	20
211	LDLQLEDNKWACNCD	20
242	IGDVVCNSPPFFKGS	20
254	KGSILSRLKKESICP	20
272	VYEEHEDPSGSLHLA	20
351	NIESLSDLRPPPQNP	20
355	LSDLRPPPQNPRKLI	20
388	TLEMLHLGNNRIEVL	20
431	FLGLHNLEYLYLEYN	20
455	FNPMPKLKVLYLNNN	20
463	VLYLNNNLLQVLPPH	20
549	PGHLDKKELKALNSE	20
612	LGLLIMFITIVFCAA	20
679	EQHMVSPMVHVYRSP	20
718	QRSLLEQENHSPLTG	20
768	QLGITEYLRKNIAQL	20
50	AKGIKMVSEISVPPS	19
56	VSEISVPPSRPFQLS	19
58	EISVPPSRPFQLSLL	19
65	RPFQLSLLNNGLTML	19
84	FSGLTNAISIHLGFN	19
100	IADIEIGAFNGLGLL	19
102	DIEIGAFNGLGLLKQ	19
108	FNGLGLLKQLHINHN	19
116	QLHINHNSLEILKED	19
162	LKVLILNDNAIESLP	19

Table XLVII-V1-HLA-DRB-0301-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
179	IFRFVPLTHLDLRGN	19
183	VPLTHLDLRGNQLQT	19
200	YVGFLEHIGRILDLQ	19
208	GRILDLQLEDNKWAC	19
226	LLQLKTWLENMPPQS	19
301	TSILKLPTKAPGLIP	19
365	PRKLILAGNIIHSLM	19
375	IHSLMKSDLVEYFTL	19
413	LQKLYLNGNHLTKLS	19
415	KLYLNGNHLTKLSKG	19
423	LTKLSKGMFLGLHNL	19
429	GMFLGLHNLEYLYLE	19
459	PKLKVLYLNNNLLQV	19
461	LKVLYLNNNLLQVLP	19
468	NNLLQVLPPHIFSGV	19
486	KVNLKTNQFTHLPVS	19
547	TSPGHLDKKELKALN	19
554	KKELKALNSEILCPG	19
604	AVPLSVLILGLLIMF	19
697	PKHLEEEEERNEKEG	19
745	EFLSFQDASSLYRNI	19
763	ERELQQLGITEYLRK	19
826	NLHAEPDYLEVLEQQ	19
13	LACISLHSQTPVLSS	18
66	PFQLSLLNNGLTMLH	18
76	LTMLHTNDFSGLTNA	18
90	AISIHLGFNNIADIE	18
164	VLILNDNAIESLPPN	18
177	PNIFRFVPLTHLDLR	18
201	VGFLEHIGRILDLQL	18
222	CNCDLLQLKTWLENM	18
287	ATSSINDSRMSTKTT	18
293	DSRMSTKTTSILKLP	18
328	YCPIPCNCKVLSPSG	18
340	PSGLLIHCQERNIES	18
341	SGLLIHCQERNIESL	18
342	GLLIHCQERNIESLS	18
367	KLILAGNIIHSLMKS	18
381	SDLVEYFTLEMLHLG	18
391	MLHLGNNRIEVLEEG	18
406	SFMNLTRLQKLYLNG	18
430	MFLGLHNLEYLYLEY	18
437	LEYLYLEYNAIKEIL	18
452	PGTFNPMPKLKVLYL	18
454	TFNPMPKLKVLYLNN	18

Table XLVII-V1-HLA-DRB-0301-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
478	IFSGVPLTKVNLKTN	18
484	LTKVNLKTNQFTHLP	18
507	DLLTQIDLEDNPWDC	18
514	LEDNPWDCSCDLVGL	18
529	QQWIQKLSKNTVTDD	18
620	TIVFCAAGIVVLVLH	18
627	GIVVLVLHRRRRYKK	18
629	VVLVLHRRRRYKKKQ	18
641	KKQVDEQMRDNSPVH	18
684	SPMVHVYRSPSFGPK	18
707	NEKEGSDAKHLQRSL	18
719	RSLLEQENHSPLTGS	18
726	NHSPLTGSNMKYKTT	18
744	TEFLSFQDASSLYRN	18
31	CDSLCNCEEKDGTML	17
96	GFNNIADIEIGAFNG	17
114	LKQLHINHNSLEILK	17
137	NLEFLQADNNFITVI	17
210	ILDLQLEDNKWACNC	17
250	PPFFKGSILSRLKKE	17
255	GSILSRLKKESICPT	17
269	TPPVYEEHEDPSGSL	17
389	LEMLHLGNNRIEVLE	17
509	LTQIDLEDNPWDCSC	17
522	SCDLVGLQQWIQKLS	17
525	LVGLQQWIQKLSKNT	17
630	VLVLHRRRRYKKKQV	17
639	YKKKQVDEQMRDNSP	17
683	VSPMVHVYRSPSFGP	17
755	LYRNILEKERELQQL	17
757	RNILEKERELQQLGI	17
788	AHYPGAHEELKLMET	17
816	TKNEYFELKANLHAE	17

Table XLVII-V3-HLA-DRB-0301-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
11	EQHMGAHEELKLMET	27

Table XLVII-V4-HLA-DRB-0301-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
5	GNIIHSLMKSILWSK	25
9	HSLMKSILWSKASGR	14
12	MKSILWSKASGRGRR	13
4	AGNIIHSLMKSILWS	12

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
81	TNDFSGLTNAISIHL	28
137	NLEFLQADNNFITVI	28
153	PSAFSKLNRLKVLIL	28
179	IFRFVPLTHLDLRGN	28
404	EGSFMNLTRLQKLYL	28
578	QTSYLMVTTPATTTN	28
2	KLWIHLFYSSLLACI	26
66	PFQLSLLNNGLTMLH	26
84	FSGLTNAISIHLGFN	26
108	FNGLGLLKQLHINHN	26
138	LEFLQADNNFITVIE	26
210	ILDLQLEDNKWACNC	26
280	SGSLHLAATSSINDS	26
388	TLEMLHLGNNRIEVL	26
398	RIEVLEEGSFMNLTR	26
437	LEYLYLEYNAIKEIL	26
460	KLKVLYLNNNLLQVL	26
503	LDDLDLLTQIDLEDN	26
522	SCDLVGLQQWIQKLS	26
554	KKELKALNSEILCPG	26
683	VSPMVHVYRSPSFGP	26
719	RSLLEQENHSPLTGS	26
1	MKLWIHLFYSSLLAC	22
6	HLFYSSLLACISLHS	22
94	HLGFNNIADIEIGAF	22
105	IGAFNGLGLLKQLHI	22
129	EDTFHGLENLEFLQA	22
144	DNNFITVIEPSAFSK	22
177	PNIFRFVPLTHLDLR	22
325	PGPYCPIPCNCKVLS	22
383	LVEYFTLEMLHLGNN	22

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
414	QKLYLNGNHLTKLSK	22
428	KGMFLGLHNLEYLYL	22
436	NLEYLYLEYNAIKEI	22
476	PHIFSGVPLTKVNLK	22
491	TNQFTHLPVSNILDD	22
615	LIMFITIVFCAAGIV	22
620	TIVFCAAGIVVLVLH	22
655	HLQYSMYGHKTTHHT	22
743	STEFLSFQDASSLYR	22
746	FLSFQDASSLYRNIL	22
787	EAHYPGAHEELKLME	22
9	YSSLLACISLHSQTP	20
10	SSLLACISLHSQTPV	20
13	LACISLHSQTPVLSS	20
43	TMLINCEAKGIKMVS	20
50	AKGIKMVSEISVPPS	20
52	GIKMVSEISVPPSRP	20
53	IKMVSEISVPPSRPF	20
73	NNGLTMLHTNDFSGL	20
90	AISIHLGFNNIADIE	20
102	DIEIGAFNGLGLLKQ	20
111	LGLLKQLHINHNSLE	20
123	SLEILKEDTFHGLEN	20
124	LEILKEDTFHGLENL	20
132	FHGLENLEFLQADNN	20
135	LENLEFLQADNNFIT	20
156	FSKLNRLKVLILNDN	20
159	LNRLKVLILNDNAIE	20
161	RLKVLILNDNAIESL	20
163	KVLILNDNAIESLPP	20
182	FVPLTHLDLRGNQLQ	20
198	LPYVGFLEHIGRILD	20
201	VGFLEHIGRILDLQL	20
204	LEHIGRILDLQLEDN	20
207	IGRILDLQLEDNKWA	20
223	NCDLLQLKTWLENMP	20
238	PQSIIGDVVCNSPPF	20
255	GSILSRLKKESICPT	20
258	LSRLKKESICPTPPV	20
269	TPPVYEEHEDPSGSL	20
300	TTSILKLPTKAPGLI	20
310	APGLIPYITKPSTQL	20
311	PGLIPYITKPSTQLP	20
340	PSGLLIHCQERNIES	20
352	IESLSDLRPPPQNPR	20

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
365	PRKLILAGNIIHSLM	20
372	GNIIHSLMKSDLVEY	20
380	KSDLVEYFTLEMLHL	20
381	SDLVEYFTLEMLHLG	20
386	YFTLEMLHLGNNRIE	20
407	FMNLTRLQKLYLNGN	20
410	LTRLQKLYLNGNHLT	20
413	LQKLYLNGNHLTKLS	20
431	FLGLHNLEYLYLEYN	20
434	LHNLEYLYLEYNAIK	20
455	FNPMPKLKVLYLNNN	20
458	MPKLKVLYLNNNLLQ	20
461	LKVLYLNNNLLQVLP	20
467	NNNLLQVLPPHIFSG	20
481	GVPLTKVNLKTNQFT	20
499	VSNILDDLDLLTQID	20
500	SNILDDLDLLTQIDL	20
506	LDLLTQIDLEDNPWD	20
509	LTQIDLEDNPWDCSC	20
525	LVGLQQWIQKLSKNT	20
529	QQWIQKLSKNTVTDD	20
537	KNTVTDDILCTSPGH	20
566	CPGLVNNPSMPTQTS	20
572	NPSMPTQTSYLMVTT	20
581	YLMVTTPATTTNTAD	20
594	ADTILRSLTDAVPLS	20
598	LRSLTDAVPLSVLIL	20
604	AVPLSVLILGLLIMF	20
606	PLSVLILGLLIMFIT	20
608	SVLILGLLIMFITIV	20
609	VLILGLLIMFITIVF	20
612	LGLLIMFITIVFCAA	20
619	ITIVFCAAGIVVLVL	20
626	AGIVVLVLHRRRRYK	20
627	GIVVLVLHRRRRYKK	20
641	KKQVDEQMRDNSPVH	20
757	RNILEKERELQQLGI	20
768	QLGITEYLRKNIAQL	20
808	PRKVLVEQTKNEYFE	20
19	HSQTPVLSSRGSCDS	18
35	CNCEEKDGTMLINCE	18
39	EKDGTMLINCEAKGI	18
65	RPFQLSLLNNGLTML	18
77	TMLHTNDFSGLTNAI	18
113	LLKQLHINHNSLEIL	18

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7		
Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
134	GLENLEFLQADNNFI	18
146	NFITVIEPSAFSKLN	18
149	TVIEPSAFSKLNRLK	18
160	NRLKVLILNDNAIES	18
183	VPLTHLDLRGNQLQT	18
215	LEDNKWACNCDLLQL	18
220	WACNCDLLQLKTWLE	18
251	PFFKGSILSRLKKES	18
252	FFKGSILSRLKKESI	18
272	VYEEHEDPSGSLHLA	18
281	GSLHLAATSSINDSR	18
287	ATSSINDSRMSTKTT	18
343	LLIHCQERNIESLSD	18
368	LILAGNIIHSLMKSD	18
369	ILAGNIIHSLMKSDL	18
488	NLKTNQFTHLPVSNI	18
514	LEDNPWDCSCDLVGL	18
553	DKKELKALNSEILCP	18
563	EILCPGLVNNPSMPT	18
564	ILCPGLVNNPSMPTQ	18
582	LMVTTPATTTNTADT	18
591	TNTADTILRSLTDAV	18
673	PSASLYEQHMVSPMV	18
698	KHLEEEEERNEKEGS	18
704	EERNEKEGSDAKHLQ	18
711	GSDAKHLQRSLLEQE	18
749	FQDASSLYRNILEKE	18
760	LEKERELQQLGITEY	18
807	RPRKVLVEQTKNEYF	18
313	LIPYITKPSTQLPGP	17
528	LQQWIQKLSKNTVTD	17
658	YSMYGHKTTHHTTER	17
818	NEYFELKANLHAEPD	17
5	IHLFYSSLLACISLH	16
197	TLPYVGFLEHIGRIL	16
200	YVGFLEHIGRILDLQ	16
217	DNKWACNCDLLQLKT	16
229	LKTWLENMPPQSIIG	16
250	PPFFKGSILSRLKKE	16
384	VEYFTLEMLHLGNNR	16
441	YLEYNAIKEILPGTF	16
452	PGTFNPMPKLKVLYL	16
462	KVLYLNNNLLQVLPP	16
675	ASLYEQHMVSPMVHV	16
687	VHVYRSPSFGPKHLE	16

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
734	NMKYKTTNQSTEFLS	16
753	SSLYRNILEKERELQ	16
802	TLMYSRPRKVLVEQT	16
817	KNEYFELKANLHAEP	16
22	TPVLSSRGSCDSLCN	15
185	LTHLDLRGNQLQTLP	15
293	DSRMSTKTTSILKLP	15
484	LTKVNLKTNQFTHLP	15
629	VVLVLHRRRRYKKKQ	15
630	VLVLHRRRRYKKKQV	15
732	GSNMKYKTTNQSTEF	15
756	YRNILEKERELQQLG	15
801	ETLMYSRPRKVLVEQ	15
15	CISLHSQTPVLSSRG	14
42	GTMLINCEAKGIKMV	14
56	VSEISVPPSRPFQLS	14
58	EISVPPSRPFQLSLL	14
68	QLSLLNNGLTMLHTN	14
69	LSLLNNGLTMLHTND	14
76	LTMLHTNDFSGLTNA	14
88	TNAISIHLGFNNIAD	14
92	SIHLGFNNIADIEIG	14
97	FNNIADIEIGAFNGL	14
100	IADIEIGAFNGLGLL	14
110	GLGLLKQLHINHNSL	14
114	LKQLHINHNSLEILK	14
116	QLHINHNSLEILKED	14
121	HNSLEILKEDTFHGL	14
145	NNFITVIEPSAFSKL	14
147	FITVIEPSAFSKLNR	14
148	ITVIEPSAFSKLNRL	14
162	LKVLILNDNAIESLP	14
164	VLILNDNAIESLPPN	14
169	DNAIESLPPNIFRFV	14
172	IESLPPNIFRFVPLT	14
176	PPNIFRFVPLTHLDL	14
187	HLDLRGNQLQTLPYV	14
192	GNQLQTLPYVGFLEH	14
195	LQTLPYVGFLEHIGR	14
208	GRILDLQLEDNKWAC	14
212	DLQLEDNKWACNCDL	14
230	KTWLENMPPQSIIGD	14
239	QSIIGDVVCNSPPFF	14
243	GDVVCNSPPFFKGSI	14
282	SLHLAATSSINDSRM	14

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
288	TSSINDSRMSTKTTS	14
314	IPYITKPSTQLPGPY	14
328	YCPIPCNCKVLSPSG	14
334	NCKVLSPSGLLIHCQ	14
341	SGLLIHCQERNIESL	14
342	GLLIHCQERNIESLS	14
349	ERNIESLSDLRPPPQ	14
355	LSDLRPPPQNPRKLI	14
366	RKLILAGNIIHSLMK	14
367	KLILAGNIIHSLMKS	14
376	HSLMKSDLVEYFTLE	14
389	LEMLHLGNNRIEVLE	14
391	MLHLGNNRIEVLEEG	14
396	NNRIEVLEEGSFMNL	14
399	IEVLEEGSFMNLTRL	14
405	GSFMNLTRLQKLYLN	14
415	KLYLNGNHLTKLSKG	14
420	GNHLTKLSKGMFLGL	14
423	LTKLSKGMFLGLHNL	14
427	SKGMFLGLHNLEYLY	14
429	GMFLGLHNLEYLYLE	14
439	YLYLEYNAIKEILPG	14
444	YNAIKEILPGTFNPM	14
447	IKEILPGTFNPMPKL	14
448	KEILPGTFNPMPKLK	14
463	VLYLNNNLLQVLPPH	14
468	NNLLQVLPPHIFSGV	14
471	LQVLPPHIFSGVPLT	14
475	PPHIFSGVPLTKVNL	14
479	FSGVPLTKVNLKTNQ	14
486	KVNLKTNQFTHLPVS	14
496	HLPVSNILDDLDLLT	14
511	QIDLEDNPWDCSCDL	14
523	CDLVGLQQWIQKLSK	14
541	TDDILCTSPGHLDKK	14
557	LKALNSEILCPGLVN	14
561	NSEILCPGLVNNPSM	14
567	PGLVNNPSMPTQTSY	14
579	TSYLMVTTPATTTNT	14
580	SYLMVTTPATTTNTA	14
595	DTILRSLTDAVPLSV	14
611	ILGLLIMFITIVFCA	14
613	GLLIMFITIVFCAAG	14
614	LLIMFITIVFCAAGI	14
616	IMFITIVFCAAGIVV	14

Table XLVIII-V1-HLA-DR1-0401-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
618	FITIVFCAAGIVVLV	14
625	AAGIVVLVLHRRRRY	14
628	IVVLVLHRRRRYKKK	14
645	DEQMRDNSPVHLQYS	14
651	NSPVHLQYSMYGHKT	14
653	PVHLQYSMYGHKTTH	14
657	QYSMYGHKTTHHTTE	14
680	QHMVSPMVHVYRSPS	14
684	SPMVHVYRSPSFGPK	14
686	MVHVYRSPSFGPKHL	14
697	PKHLEEEEERNEKEG	14
718	QRSLLEQENHSPLTG	14
727	HSPLTGSNMKYKTTN	14
744	TEFLSFQDASSLYRN	14
763	ERELQQLGITEYLRK	14
766	LQQLGITEYLRKNIA	14
772	TEYLRKNIAQLQPDM	14
776	RKNIAQLQPDMEAHY	14
779	IAQLQPDMEAHYPGA	14
794	HEELKLMETLMYSRP	14
797	LKLMETLMYSRPRKV	14
800	METLMYSRPRKVLVE	14
810	KVLVEQTKNEYFELK	14
820	YFELKANLHAEPDYL	14
824	KANLHAEPDYLEVLE	14

Table XLVIII-V3-HLA-DR1-0401-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
5	PSASLYEQHMGAHEE	18
11	EQHMGAHEELKLMET	14
1	TTERPSASLYEQHMG	12
3	ERPSASLYEQHMGAH	12
9	LYEQHMGAHEELKLM	12
10	YEQHMGAHEELKLME	12
12	QHMGAHEELKLMETL	12
14	MGAHEELKLMETLMY	12
7	ASLYEQHMGAHEELK	10
6	SASLYEQHMGAHEEL	8

Table XLVIII-V4-HLA-DR1-0401-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
5	GNIIHSLMKSILWSK	20
9	HSLMKSILWSKASGR	20
1	LILAGNIIHSLMKSI	18
2	ILAGNIIHSLMKSIL	18
10	SLMKSILWSKASGRG	18
14	SILWSKASGRGRREE	16
4	AGNIIHSLMKSILWS	14
8	IHSLMKSILWSKASG	14
13	KSILWSKASGRGRRE	9

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
683	VSPMVHVYRSPSFGP	26
153	PSAFSKLNRLKVLIL	25
452	PGTFNPMPKLKVLYL	25
179	IFRFVPLTHLDLRGN	24
404	EGSFMNLTRLQKLYL	24
615	LIMFITIVFCAAGIV	24
627	GIVVLVLHRRRRYKK	24
6	HLFYSSLLACISLHS	23
81	TNDFSGLTNAISIHL	23
441	YLEYNAIKEILPGTF	23
626	AGIVVLVLHRRRRYK	23
144	DNNFITVIEPSAFSK	22
407	FMNLTRLQKLYLNGN	22
420	GNHLTKLSKGMFLGL	22
680	QHMVSPMVHVYRSPS	22
173	ESLPPNIFRFVPLTH	21
201	VGFLEHIGRILDLQL	21
328	YCPIPCNCKVLSPSG	21
769	LGITEYLRKNIAQLQ	21
198	LPYVGFLEHIGRILD	20
239	QSIIGDVVCNSPPFF	20
254	KGSILSRLKKESICP	20
255	GSILSRLKKESICPT	20
301	TSILKLPTKAPGLIP	20
311	PGLIPYITKPSTQLP	20

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
349	ERNIESLSDLRPPPQ	20
372	GNIIHSLMKSDLVEY	20
529	QQWIQKLSKNTVTDD	20
616	IMFITIVFCAAGIVV	20
641	KKQVDEQMRDNSPVH	20
797	LKLMETLMYSRPRKV	20
820	YFELKANLHAEPDYL	20
384	VEYFTLEMLHLGNNR	19
595	DTILRSLTDAVPLSV	19
802	TLMYSRPRKVLVEQT	19
53	IKMVSEISVPPSRPF	18
132	FHGLENLEFLQADNN	18
300	TTSILKLPTKAPGLI	18
414	QKLYLNGNHLTKLSK	18
491	TNQFTHLPVSNILDD	18
655	HLQYSMYGHKTTHHT	18
817	KNEYFELKANLHAEP	18
105	IGAFNGLGLLKQLHI	17
197	TLPYVGFLEHIGRIL	17
383	LVEYFTLEMLHLGNN	17
476	PHIFSGVPLTKVNLK	17
516	DNPWDCSCDLVGLQQ	17
625	AAGIVVLVLHRRRRY	17
628	IVVLVLHRRRRYKKK	17
1	MKLWIHLFYSSLLAC	16
18	LHSQTPVLSSRGSCD	16
64	SRPFQLSLLNNGLTM	16
94	HLGFNNIADIEIGAF	16
129	EDTFHGLENLEFLQA	16
177	PNIFRFVPLTHLDLR	16
229	LKTWLENMPPQSIIG	16
270	PPVYEEHEDPSGSLH	16
297	STKTTSILKLPTKAP	16
325	PGPYCPIPCNCKVLS	16
427	SKGMFLGLHNLEYLY	16
428	KGMFLGLHNLEYLYL	16
436	NLEYLYLEYNAIKEI	16
578	QTSYLMVTTPATTTN	16
743	STEFLSFQDASSLYR	16
753	SSLYRNILEKERELQ	16
754	SLYRNILEKERELQQ	16
768	QLGITEYLRKNIAQL	16
818	NEYFELKANLHAEPD	16
43	TMLINCEAKGIKMVS	15
46	INCEAKGIKMVSEIS	15

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
49	EAKGIKMVSEISVPP	15
107	AFNGLGLLKQLHINH	15
145	NNFITVIEPSAFSKL	15
182	FVPLTHLDLRGNQLQ	15
252	FFKGSILSRLKKESI	15
314	IPYITKPSTQLPGPY	15
342	GLLIHCQERNIESLS	15
368	LILAGNIIHSLMKSD	15
591	TNTADTILRSLTDAV	15
602	TDAVPLSVLILGLLI	15
629	VVLVLHRRRRYKKKQ	15
630	VLVLHRRRRYKKKQV	15
673	PSASLYEQHMVSPMV	15
711	GSDAKHLQRSLLEQE	15
749	FQDASSLYRNILEKE	15
756	YRNILEKERELQQLG	15
801	ETLMYSRPRKVLVEQ	15
19	HSQTPVLSSRGSCDS	14
39	EKDGTMLINCEAKGI	14
55	MVSEISVPPSRPFQL	14
72	LNNGLTMLHTNDFSG	14
73	NNGLTMLHTNDFSGL	14
110	GLGLLKQLHINHNSL	14
113	LLKQLHINHNSLEIL	14
120	NHNSLEILKEDTFHG	14
227	LQLKTWLENMPPQSI	14
238	PQSIIGDVVCNSPPF	14
268	PTPPVYEEHEDPSGS	14
276	HEDPSGSLHLAATSS	14
291	INDSRMSTKTTSILK	14
338	LSPSGLLIHCQERNI	14
351	NIESLSDLRPPPQNP	14
385	EYFTLEMLHLGNNRI	14
388	TLEMLHLGNNRIEVL	14
417	YLNGNHLTKLSKGMF	14
468	NNLLQVLPPHIFSGV	14
469	NLLQVLPPHIFSGVP	14
478	IFSGVPLTKVNLKTN	14
481	GVPLTKVNLKTNQFT	14
506	LDLLTQIDLEDNPWD	14
526	VGLQQWIQKLSKNTV	14
537	KNTVTDDILCTSPGH	14
546	CTSPGHLDKKELKAL	14
563	EILCPGLVNNPSMPT	14
577	TQTSYLMVTTPATTT	14

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
604	AVPLSVLILGLLIMF	14
664	KTTHHTTERPSASLY	14
681	HMVSPMVHVYRSPSF	14
701	EEEEERNEKEGSDAK	14
719	RSLLEQENHSPLTGS	14
781	QLQPDMEAHYPGAHE	14
809	RKVLVEQTKNEYFEL	14
15	CISLHSQTPVLSSRG	13
41	DGTMLINCEAKGIKM	13
66	PFQLSLLNNGLTMLH	13
85	SGLTNAISIHLGFNN	13
90	AISIHLGFNNIADIE	13
156	FSKLNRLKVLILNDN	13
159	LNRLKVLILNDNAIE	13
169	DNAIESLPPNIFRFV	13
223	NCDLLQLKTWLENMP	13
240	SIIGDVVCNSPPFFK	13
321	STQLPGPYCPIPCNC	13
396	NNRIEVLEEGSFMNL	13
458	MPKLKVLYLNNNLLQ	13
460	KLKVLYLNNNLLQVL	13
464	LYLNNNLLQVLPPHI	13
472	QVLPPHIFSGVPLTK	13
496	HLPVSNILDDLDLLT	13
522	SCDLVGLQQWIQKLS	13
525	LVGLQQWIQKLSKNT	13
554	KKELKALNSEILCPG	13
606	PLSVLILGLLIMFIT	13
609	VLILGLLIMFITIVF	13
611	ILGLLIMFITIVFCA	13
614	LLIMFITIVFCAAGI	13
9	YSSLLACISLHSQTP	12
10	SSLLACISLHSQTPV	12
12	LLACISLHSQTPVLS	12
22	TPVLSSRGSCDSLCN	12
31	CDSLCNCEEKDGTML	12
50	AKGIKMVSEISVPPS	12
52	GIKMVSEISVPPSRP	12
75	GLTMLHTNDFSGLTN	12
97	FNNIADIEIGAFNGL	12
99	NIADIEIGAFNGLGL	12
108	FNGLGLLKQLHINHN	12
111	LGLLKQLHINHNSLE	12
121	HNSLEILKEDTFHGL	12
123	SLEILKEDTFHGLEN	12

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
135	LENLEFLQADNNFIT	12
142	QADNNFITVIEPSAF	12
160	NRLKVLILNDNAIES	12
161	RLKVLILNDNAIESL	12
163	KVLILNDNAIESLPP	12
166	ILNDNAIESLPPNIF	12
192	GNQLQTLPYVGFLEH	12
195	LQTLPYVGFLEHIGR	12
200	YVGFLEHIGRILDLQ	12
204	LEHIGRILDLQLEDN	12
207	IGRILDLQLEDNKWA	12
210	ILDLQLEDNKWACNC	12
226	LLQLKTWLENMPPQS	12
230	KTWLENMPPQSIIGD	12
250	PPFFKGSILSRLKKE	12
260	RLKKESICPTPPVYE	12
269	TPPVYEEHEDPSGSL	12
279	PSGSLHLAATSSIND	12
310	APGLIPYITKPSTQL	12
331	IPCNCKVLSPSGLLI	12
352	IESLSDLRPPPQNPR	12
366	RKLILAGNIIHSLMK	12
386	YFTLEMLHLGNNRIE	12
395	GNNRIEVLEEGSFMN	12
410	LTRLQKLYLNGNHLT	12
431	FLGLHNLEYLYLEYN	12
434	LHNLEYLYLEYNAIK	12
444	YNAIKEILPGTFNPM	12
448	KEILPGTFNPMPKLK	12
455	FNPMPKLKVLYLNNN	12
465	YLNNNLLQVLPPHIF	12
467	NNNLLQVLPPHIFSG	12
470	LLQVLPPHIFSGVPL	12
500	SNILDDLDLLTQIDL	12
503	LDDLDLLTQIDLEDN	12
511	QIDLEDNPWDCSCDL	12
538	NTVTDDILCTSPGHL	12
539	TVTDDILCTSPGHLD	12
551	HLDKKELKALNSEIL	12
557	LKALNSEILCPGLVN	12
562	SEILCPGLVNNPSMP	12
569	LVNNPSMPTQTSYLM	12
576	PTQTSYLMVTTPATT	12
608	SVLILGLLIMFITIV	12
613	GLLIMFITIVFCAAG	12

Table XLIX-V1-HLA-DRB1-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 3; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
642	KQVDEQMRDNSPVHL	12
648	MRDNSPVHLQYSMYG	12
651	NSPVHLQYSMYGHKT	12
674	SASLYEQHMVSPMVH	12
686	MVHVYRSPSFGPKHL	12
718	QRSLLEQENHSPLTG	12
732	GSNMKYKTTNQSTEF	12
741	NQSTEFLSFQDASSL	12
763	ERELQQLGITEYLRK	12
773	EYLRKNIAQLQPDME	12
776	RKNIAQLQPDMEAHY	12
780	AQLQPDMEAHYPGAH	12
794	HEELKLMETLMYSRP	12

Table XLIX-V3-HLA-DRB-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 7; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
5	PSASLYEQHMGAHEE	14
7	ASLYEQHMGAHEELK	10
9	LYEQHMGAHEELKLM	8
13	HMGAHEELKLMETLM	8
11	EQHMGAHEELKLMET	7
3	ERPSASLYEQHMGAH	6
4	RPSASLYEQHMGAHE	6
6	SASLYEQHMGAHEEL	6
8	SLYEQHMGAHEELKL	6
14	MGAHEELKLMETLMY	6

Table XLIX-V4-HLA-DRB-1101-15mers-158P1D7

Each peptide is a portion of SEQ ID NO: 9; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
5	GNIIHSLMKSILWSK	21
9	HSLMKSILWSKASGR	19

Table XLIX-V4-HLA-DRB-1101-15mers-158P1D7		
1	LILAGNIIHSLMKSI	15
11	LMKSILWSKASGRGR	14
13	KSILWSKASGRGRRE	14
10	SLMKSILWSKASGRG	12
14	SILWSKASGRGRREE	11

Table XXII – 158P1D7 v.6 – HLA-A1-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
11	SFGPKHLEE	12
10	PSFGPKHLE	8
1	GNIIHSLMN	7
7	LMNPSFGPK	7

Table XXIII – 158P1D7 v.6 – HLA-A0201-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
6	SLMNPSFGP	15
2	NIIHSLMNP	14
7	LMNPSFGPK	13
3	IIHSLMNPS	12
9	NPSFGPKHL	10
11	SFGPKHLEE	8

Table XXIV – 158P1D7 v.6 – HLA-A0203-9-mers

Pos	123456789	score
	No results found	

Table XXV – 158P1D7 v.6 – HLA-A3-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
7	LMNPSFGPK	14
2	NIIHSLMNP	12
6	SLMNPSFGP	12
3	IIHSLMNPS	10
15	KHLEEEEER	10
4	IHSLMNPSF	9
1	GNIIHSLMN	8
11	SFGPKHLEE	8
8	MNPSFGPKH	7

Table XXVI – 158P1D7 v.6 – HLA-A26-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
2	NIIHSLMNP	12
4	IHSLMNPSF	9
9	NPSFGPKHL	8
1	GNIIHSLMN	6
3	IIHSLMNPS	6

Table XXVII – 158P1D7 v.6 – HLA-B0702-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
9	NPSFGPKHL	22
4	IHSLMNPSF	10
13	GPKHLEEEE	10

Table XXVIII – 158P1D7 v.6 – HLA-B08-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
13	GPKHLEEEE	18
9	NPSFGPKHL	17
11	SFGPKHLEE	13
4	IHSLMNPSF	9
6	SLMNPSFGP	8

Table XXIX – 158P1D7 v.6 – HLA-B1510-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
4	IHSLMNPSF	20
9	NPSFGPKHL	13
15	KHLEEEEER	12

Table XXX – 158P1D7 v.6 – HLA-B2705-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
15	KHLEEEEER	17
4	IHSLMNPSF	15
7	LMNPSFGPK	12
9	NPSFGPKHL	12
8	MNPSFGPKH	11

Table XXXI – 158P1D7 v.6 – HLA-B2709-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
4	IHSLMNPSF	10
9	NPSFGPKHL	10
1	GNIIHSLMN	5

Table XXXII – 158P1D7 v.6 – HLA-B4402-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
9	NPSFGPKHL	15
4	IHSLMNPSF	12

Table XXXIII – 158P1D7 v.6 – HLA-B5101-9-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	123456789	score
9	NPSFGPKHL	20
12	FGPKHLEEE	10
13	GPKHLEEEE	10

Table XXXIV – 158P1D7 v.6 – HLA-A1-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 9 amino acids, and the end position for each peptide is the start position plus eight.

Pos	1234567890	score
11	PSFGPKHLEE	11
1	AGNIIHSLMN	8
8	LMNPSFGPKH	7
7	SLMNPSFGPK	6
12	SFGPKHLEEE	6

Table XXXV – 158P1D7 v.6 – HLA-A0201-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
8	LMNPSFGPKH	16
4	IIHSLMNPSF	13
7	SLMNPSFGPK	13
3	NIIHSLMNPS	11
12	SFGPKHLEEE	10
9	MNPSFGPKHL	9

Table XXXVI – 158P1D7 v.6 – HLA-A0203-10-mers

Pos	1234567890	score
	No Results found	

Table XXXVII – 158P1D7 v.6 – HLA-A3-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
7	SLMNPSFGPK	23
4	IIHSLMNPSF	16
3	NIIHSLMNPS	11
8	LMNPSFGPKH	10

Table XXXVIII – 158P1D7 v.6 – HLA-A26-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
4	IIHSLMNPSF	14
9	MNPSFGPKHL	9
2	GNIIHSLMNP	8
3	NIIHSLMNPS	8
12	SFGPKHLEEE	6

Table XXXIX – 158P1D7 v.6 – HLA-A0702-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score

Table XXXIX – 158P1D7 v.6 – HLA-A0702-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
10	NPSFGPKHLE	12
9	MNPSFGPKHL	10
14	GPKHLEEEEE	10
4	IIHSLMNPSF	8

Table XL – 158P1D7 v.6 – HLA-B08-10-mers

Pos	1234567890	score
	No Results Found	

Table XLI – 158P1D7 v.6 – HLA-B1510-10-mers

Pos	1234567890	score
	No Results Found	

Table XLII – 158P1D7 v.6 – HLA-B2705-10-mers

Pos	1234567890	score
	No Results Found	

Table XLIII – 158P1D7 v.6 – HLA-B2709-10-mers

Pos	1234567890	score
	No Results Found	

Table XLIV – 158P1D7 v.6 – HLA-B4402-10-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 10 amino acids, and the end position for each peptide is the start position plus nine.

Pos	1234567890	score
9	MNPSFGPKHL	14
4	IIHSLMNPSF	10

Table XLV – 158P1D7 v.6 – HLA-B5101-10-mers

Pos	1234567890	score
	No Results Found	

Table XLVI – 158P1D7 v.6 – HLA-DRB 0101-15-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
7	GNIIHSLMNPSFGPK	24
11	HSLMNPSFGPKHLEE	18
9	IIHSLMNPSFGPKHL	17
1	RKLILAGNIIHSLMN	16
2	KLILAGNIIHSLMNP	16
4	ILAGNIIHSLMNPSF	16
6	AGNIIHSLMNPSFGP	16
12	SLMNPSFGPKHLEEE	16
3	LILAGNIIHSLMNPS	14
8	NIIHSLMNPSFGPKH	13

Table XLVII – 158P1D7 v.6 – HLA-DRB -0301-15-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
3	LILAGNIIHSLMNPS	15
1	RKLILAGNIIHSLMN	12
6	AGNIIHSLMNPSFGP	12

Table XLVII – 158P1D7 v.6 – HLA-DRB -0301-15-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
7	GNIIHSLMNPSFGPK	25
2	KLILAGNIIHSLMNP	18
6	AGNIIHSLMNPSFGP	13
1	RKLILAGNIIHSLMN	12
11	HSLMNPSFGPKHLEE	12

Table XLVIII – 158P1D7 v.6 – HLA-DRB 0410-15-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
7	GNIIHSLMNPSFGPK	20
3	LILAGNIIHSLMNPS	18
4	ILAGNIIHSLMNPSF	18
1	RKLILAGNIIHSLMN	14
2	KLILAGNIIHSLMNP	14
6	AGNIIHSLMNPSFGP	14
10	IHSLMNPSFGPKHLE	14
12	SLMNPSFGPKHLEEE	12

Table XLIX – 158P1D7 v.6 – HLA-DRB 1101-15-mers

Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.

Pos	123456789012345	score
3	LILAGNIIHSLMNPS	15
1	RKLILAGNIIHSLMN	12
6	AGNIIHSLMNPSFGP	12

Table XLIX – 158P1D7 v.6 – HLA-DRB 1101-15-mers		
Each peptide is a portion of SEQ ID NO: 13; each start position is specified, the length of peptide is 15 amino acids, and the end position for each peptide is the start position plus fourteen.		
Pos	123456789012345	score
7	GNIIHSLMNPSFGPK	12
8	NIIHSLMNPSFGPKH	12
15	NPSFGPKHLEEEEER	10
13	LMNPSFGPKHLEEEE	9
14	MNPSFGPKHLEEEEE	9
11	HSLMNPSFGPKHLEE	7

Table L. Exon boundaries of transcript 158P1D7 v.1

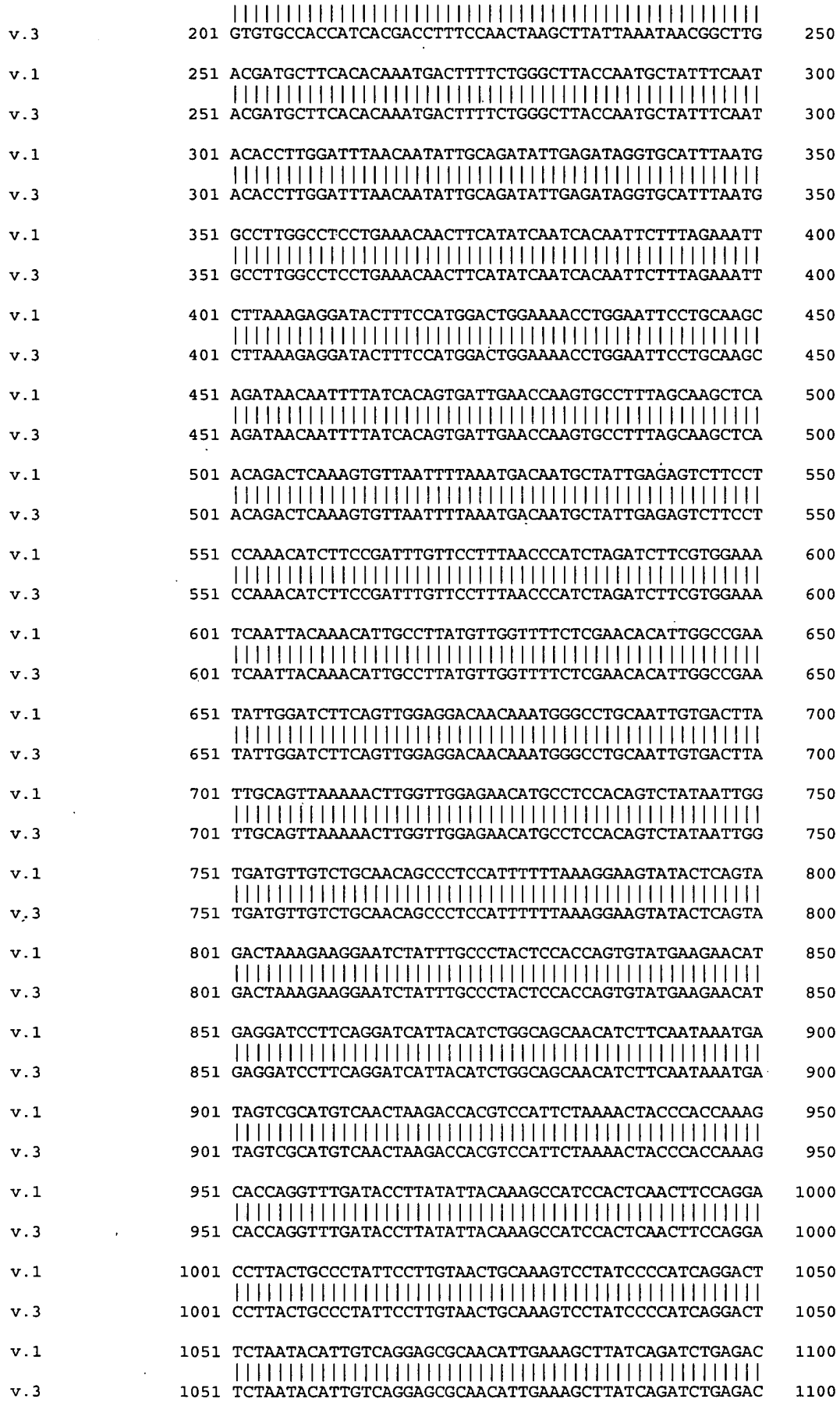
Exon	Start	End	Length
1	1	2555	2555

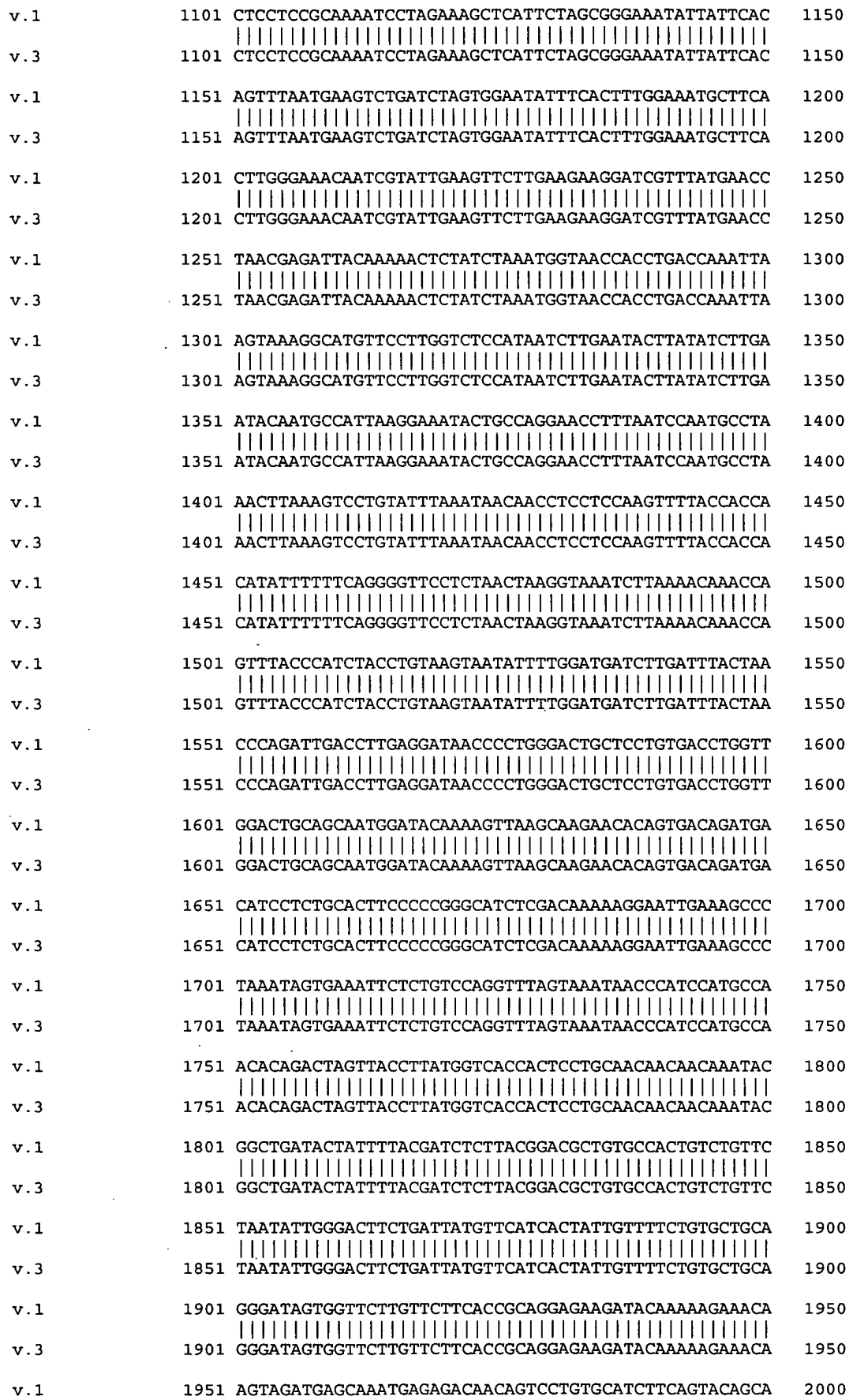
Table LI(a). Nucleotide sequence of transcript variant 158P1D7 v.3 (SEQ ID NO: 70)

```
tcggatttca tcacatgaca acatgaagct gtggattcat ctcttttatt catctctcct      60
tgcctgtata tctttacact cccaaactcc agtgctctca tccagaggct cttgtgattc     120
tctttgcaat tgtgaggaaa aagatggcac aatgctaata aattgtgaag caaaaggtat     180
caagatggta tctgaaataa gtgtgccacc atcacgacct ttccaactaa gcttattaaa     240
taacggcttg acgatgcttc acacaaatga cttttctggg cttaccaatg ctatttcaat     300
acaccttgga tttaacaata ttgcagatat tgagataggt gcatttaatg gccttggcct     360
cctgaaacaa cttcatatca atcacaattc tttagaaatt cttaaagagg atactttcca     420
tggactggaa aacctggaat tcctgcaagc agataacaat tttatcacag tgattgaacc     480
aagtgccttt agcaagctca acagactcaa agtgttaatt ttaaatgaca atgctattga     540
gagtcttcct ccaaacatct tccgatttgt tcctttaacc catctagatc ttcgtggaaa     600
tcaattacaa acattgcctt atgttggttt tctcgaacac attggccgaa tattggatct     660
tcagttggag gacaacaaat gggcctgcaa ttgtgactta ttgcagttaa aaacttggtt     720
ggagaacatg cctccacagt ctataattgg tgatgttgtc tgcaacagcc ctccattttt     780
taaaggaagt atactcagta gactaaagaa ggaatctatt tgccctactc caccagtgta     840
tgaagaacat gaggatcctt caggatcatt acatctggca gcaacatctt caataaatga     900
tagtcgcatg tcaactaaga ccacgtccat tctaaaacta cccaccaaag caccaggttt     960
gataccttat attacaaagc catccactca acttccagga ccttactgcc ctattccttg    1020
taactgcaaa gtcctatccc catcaggact tctaatacat tgtcaggagc gcaacattga    1080
aagcttatca gatctgagac ctcctccgca aaatcctaga aagctcattc tagcgggaaa    1140
tattattcac agtttaatga agtctgatct agtggaatat ttcactttgg aaatgcttca    1200
cttgggaaac aatcgtattg aagttcttga agaaggatcg tttatgaacc taacgagatt    1260
acaaaaactc tatctaaatg gtaaccacct gaccaaatta agtaaaggca tgttccttgg    1320
tctccataat cttgaatact tatatcttga atacaatgcc attaaggaaa tactgccagg    1380
aacctttaat ccaatgccta aacttaaagt cctgtattta aataacaacc tcctccaagt    1440
tttaccacca catatttttt caggggttcc tctaactaag gtaaatctta aaacaaacca    1500
gtttacccat ctacctgtaa gtaatatttt ggatgatctt gatttactaa cccagattga    1560
ccttgaggat aacccctggg actgctcctg tgacctggtt ggactgcagc aatggataca    1620
aaagttaagc aagaacacag tgacagatga catcctctgc acttcccccg ggcatctcga    1680
caaaaaggaa ttgaaagccc taaatagtga aattctctgt ccaggtttag taaataaccc    1740
atccatgcca acacagacta gttaccttat ggtcaccact cctgcaacaa caacaaatac    1800
ggctgatact attttacgat ctcttacgga cgctgtgcca ctgtctgttc taatattggg    1860
acttctgatt atgttcatca ctattgtttt ctgtgctgca gggatagtgg ttcttgttct    1920
tcaccgcagg agaagataca aaaagaaaca agtagatgag caaatgagag acaacagtcc    1980
tgtgcatctt cagtacagca tgtatggcca taaaaccact catcacacta ctgaaagacc    2040
ctctgcctca ctctatgaac agcacatggg agcccacgaa gagctgaagt taatggaaac    2100
attaatgtac tcacgtccaa ggaaggtatt agtggaacag acaaaaaatg agtattttga    2160
acttaaagct aatttacatg ctgaacctga ctatttagaa gtcctggagc agcaaacata    2220
gatggaga                                                            2228
```

Table LII(a). Nucleotide sequence alignment of 158P1D7 v.1 (SEQ ID NO: 71) and 158P1D7 v.3 (SEQ ID NO: 72)

```
v.1          1 TCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTCTTTTATT     50
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          1 TCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTCTTTTATT     50

v.1         51 CATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGTGCTCTCA    100
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3         51 CATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGTGCTCTCA    100

v.1        101 TCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAGATGGCAC    150
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        101 TCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAGATGGCAC    150

v.1        151 AATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCTGAAATAA    200
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        151 AATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCTGAAATAA    200

v.1        201 GTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAACGGCTTG    250
```


v.3 201 GTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAACGGCTTG 250
v.1 251 ACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT 300
v.3 251 ACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT 300
v.1 301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATG 350
v.3 301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATG 350
v.1 351 GCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATT 400
v.3 351 GCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATT 400
v.1 401 CTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCCTGCAAGC 450
v.3 401 CTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCCTGCAAGC 450
v.1 451 AGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGCAAGCTCA 500
v.3 451 AGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGCAAGCTCA 500
v.1 501 ACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAGTCTTCCT 550
v.3 501 ACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAGTCTTCCT 550
v.1 551 CCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTCGTGGAAA 600
v.3 551 CCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTCGTGGAAA 600
v.1 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAA 650
v.3 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAA 650
v.1 651 TATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTA 700
v.3 651 TATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTA 700
v.1 701 TTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTATAATTGG 750
v.3 701 TTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTATAATTGG 750
v.1 751 TGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATACTCAGTA 800
v.3 751 TGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATACTCAGTA 800
v.1 801 GACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGAAGAACAT 850
v.3 801 GACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGAAGAACAT 850
v.1 851 GAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA 900
v.3 851 GAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA 900
v.1 901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAG 950
v.3 901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAG 950
v.1 951 CACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGA 1000
v.3 951 CACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGA 1000
v.1 1001 CCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCATCAGGACT 1050
v.3 1001 CCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCATCAGGACT 1050
v.1 1051 TCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGATCTGAGAC 1100
v.3 1051 TCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGATCTGAGAC 1100


v.1 1101 CTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATATTATTCAC 1150
v.3 1101 CTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATATTATTCAC 1150
v.1 1151 AGTTTAATGAAGTCTGATCTAGTGGAATATTTCACTTTGGAAATGCTTCA 1200
v.3 1151 AGTTTAATGAAGTCTGATCTAGTGGAATATTTCACTTTGGAAATGCTTCA 1200
v.1 1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACC 1250
v.3 1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACC 1250
v.1 1251 TAACGAGATTACAAAAACTCTATCTAAATGGTAACCACCTGACCAAATTA 1300
v.3 1251 TAACGAGATTACAAAAACTCTATCTAAATGGTAACCACCTGACCAAATTA 1300
v.1 1301 AGTAAAGGCATGTTCCTTGGTCTCCATAATCTTGAATACTTATATCTTGA 1350
v.3 1301 AGTAAAGGCATGTTCCTTGGTCTCCATAATCTTGAATACTTATATCTTGA 1350
v.1 1351 ATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCAATGCCTA 1400
v.3 1351 ATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCAATGCCTA 1400
v.1 1401 AACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTTACCACCA 1450
v.3 1401 AACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTTACCACCA 1450
v.1 1451 CATATTTTTTCAGGGGTTCCTCTAACTAAGGTAAATCTTAAAACAAACCA 1500
v.3 1451 CATATTTTTTCAGGGGTTCCTCTAACTAAGGTAAATCTTAAAACAAACCA 1500
v.1 1501 GTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGATTTACTAA 1550
v.3 1501 GTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGATTTACTAA 1550
v.1 1551 CCCAGATTGACCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTT 1600
v.3 1551 CCCAGATTGACCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTT 1600
v.1 1601 GGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGACAGATGA 1650
v.3 1601 GGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGACAGATGA 1650
v.1 1651 CATCCTCTGCACTTCCCCCGGGCATCTCGACAAAAAGGAATTGAAAGCCC 1700
v.3 1651 CATCCTCTGCACTTCCCCCGGGCATCTCGACAAAAAGGAATTGAAAGCCC 1700
v.1 1701 TAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATCCATGCCA 1750
v.3 1701 TAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATCCATGCCA 1750
v.1 1751 ACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC 1800
v.3 1751 ACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC 1800
v.1 1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTC 1850
v.3 1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTC 1850
v.1 1851 TAATATTGGGACTTCTGATTATGTTCATCACTATTGTTTTCTGTGCTGCA 1900
v.3 1851 TAATATTGGGACTTCTGATTATGTTCATCACTATTGTTTTCTGTGCTGCA 1900
v.1 1901 GGGATAGTGGTTCTTGTTCTTCACCGCAGGAGAAGATACAAAAAGAAACA 1950
v.3 1901 GGGATAGTGGTTCTTGTTCTTCACCGCAGGAGAAGATACAAAAAGAAACA 1950
v.1 1951 AGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAGTACAGCA 2000

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          1951 AGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAGTACAGCA   2000

v.1          2001 TGTATGGCCATAAAACCACTCATCACACTACTGAAAGACCCTCTGCCTCA   2050
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          2001 TGTATGGCCATAAAACCACTCATCACACTACTGAAAGACCCTCTGCCTCA   2050

v.1          2051 CTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATAGAAGTCC   2100
                  ||||||||||||||||||
v.3          2051 CTCTATGAACAGCACATGG----------------------------   2069

v.1          2101 ATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAATGAGAAAG   2150

v.3          2070 ----------------------------------------------   2069

v.1          2151 AAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACAGGAAAAT   2200

v.3          2070 ----------------------------------------------   2069

v.1          2201 CATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGAACCAATC   2250

v.3          2070 ----------------------------------------------   2069

v.1          2251 AACAGAATTTTTATCCTTCCAAGATGCCAGCTCATTGTACAGAAACATTT   2300

v.3          2070 ----------------------------------------------   2069

v.1          2301 TAGAAAAAGAAAGGGAACTTCAGCAACTGGGAATCACAGAATACCTAAGG   2350

v.3          2070 ----------------------------------------------   2069

v.1          2351 AAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGC   2400
                                                              ||||
v.3          2070 ----------------------------------------------GAGC   2073

v.1          2401 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGA   2450
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          2074 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGA   2123

v.1          2451 AGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAAAGCTAAT   2500
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          2124 AGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAAAGCTAAT   2173

v.1          2501 TTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATAGAT   2550
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3          2174 TTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATAGAT   2223

v.1          2551 GGAGA   2555
                  |||||
v.3          2224 GGAGA   2228
```

Table LIII(a). Peptide sequences of protein coded by 158P1D7 v.3 (SEQ ID NO: 73)

```
MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLCNCEEK DGTMLINCEA KGIKMVSEIS      60
VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIHLGFNNI ADIEIGAFNG LGLLKQLHIN     120
HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF     180
RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS     240
IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT     300
TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP     360
PPQNPRKLIL AGNIIHSLMK SDLVEYFTLE MLHLGNNRIE VLEEGSFMNL TRLQKLYLNG     420
NHLTKLSKGM FLGLHNLEYL YLEYNAIKEI LPGTFNPMPK LKVLYLNNNL LQVLPPHIFS     480
GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIQKLSKNTV     540
TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS     600
LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM     660
YGHKTTHHTT ERPSASLYEQ HMGAHEELKL METLMYSRPR KVLVEQTKNE YFELKANLHA     720
EPDYLEVLEQ QT                                                         732
```

Table LIV(a). Amino acid sequence alignment of 158P1D7 v.1 (SEQ ID NO: 74) and 158P1D7 v.3 (SEQ ID NO: 75)

```
v.1            1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
v.3          1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50

v.1         51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3         51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100

v.1        101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150

v.1        151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200

v.1        201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250

v.1        251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300

v.1        301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350

v.1        351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400

v.1        401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450

v.1        451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500

v.1        501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550

v.1        551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600

v.1        601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.3        601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650

v.1        651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHL    700
               ||||||||||||||||||||||||||||||
v.3        651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHM----------------     682

v.1        701 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    750

v.3        683 -------------------------------------------------     682

v.1        751 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    800
                                                   |||||||||
v.3        683 ----------------------------------------GAHEELKLM    691

v.1        801 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    841
               ||||||||||||||||||||||||||||||||||||||||
v.3        692 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    732
```

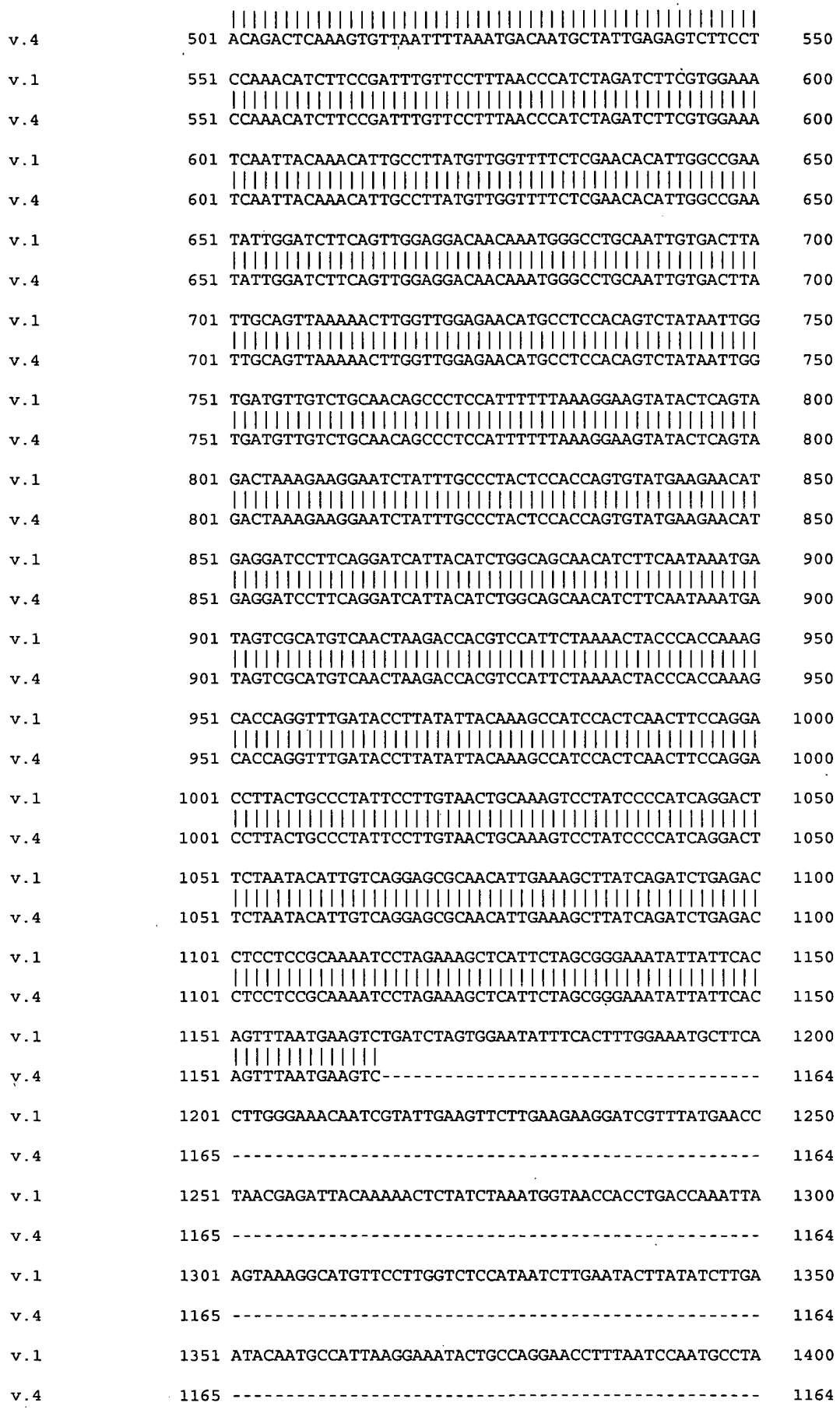
Table LI(b). Nucleotide sequence of transcript variant 158P1D7 v.4 (SEQ ID NO: 76)

```
tcggatttca tcacatgaca acatgaagct gtggattcat ctcttttatt catctctcct      60
tgcctgtata tctttacact cccaaactcc agtgctctca tccagaggct cttgtgattc     120
```

```
tctttgcaat tgtgaggaaa aagatggcac aatgctaata aattgtgaag caaaaggtat     180
caagatggta tctgaaataa gtgtgccacc atcacgacct ttccaactaa gcttattaaa     240
taacggcttg acgatgcttc acacaaatga cttttctggg cttaccaatg ctatttcaat     300
acaccttgga tttaacaata ttgcagatat tgagataggt gcatttaatg gccttggcct     360
cctgaaacaa cttcatatca atcacaattc tttagaaatt cttaaagagg atactttcca     420
tggactggaa aacctggaat tcctgcaagc agataacaat tttatcacag tgattgaacc     480
aagtgccttt agcaagctca acagactcaa agtgttaatt ttaaatgaca atgctattga     540
gagtcttcct ccaaacatct tccgatttgt tcctttaacc catctagatc ttcgtggaaa     600
tcaattacaa acattgcctt atgttggttt tctcgaacac attggccgaa tattggatct     660
tcagttggag gacaacaaat gggcctgcaa ttgtgactta ttgcagttaa aaacttggtt     720
ggagaacatg cctccacagt ctataattgg tgatgttgtc tgcaacagcc ctccattttt     780
taaaggaagt atactcagta gactaaagaa ggaatctatt tgccctactc caccagtgta     840
tgaagaacat gaggatcctt caggatcatt acatctggca gcaacatctt caataaatga     900
tagtcgcatg tcaactaaga ccacgtccat tctaaaacta cccaccaaag caccaggttt     960
gataccttat attacaaagc catccactca acttccagga ccttactgcc ctattccttg    1020
taactgcaaa gtcctatccc catcaggact tctaatacat tgtcaggagc gcaacattga    1080
aagcttatca gatctgagac ctcctccgca aaatcctaga aagctcattc tagcgggaaa    1140
tattattcac agtttaatga agtccatcct ttggtccaaa gcatctggaa gaggaagaag    1200
agaggaatga gaaagaagga agtgatgcaa aacatctcca aagaagtctt ttggaacagg    1260
aaaatcattc accactcaca gggtcaaata tgaaatacaa aaccacgaac caatcaacag    1320
aatttttatc cttccaagat gccagctcat tgtacagaaa cattttagaa aaagaaaggg    1380
aacttcagca actgggaatc acagaatacc taaggaaaaa cattgctcag ctccagcctg    1440
atatggaggc acattatcct ggagcccacg aagagctgaa gttaatggaa acattaatgt    1500
actcacgtcc aaggaaggta ttagtggaac agacaaaaaa tgagtatttt gaacttaaag    1560
ctaatttaca tgctgaacct gactatttag aagtcctgga gcagcaaaca tagatggaga    1620
```

Table LII(b). Nucleotide sequence alignment of 158P1D7 v.1 (SEQ ID NO: 77) and 158P1D7 v.4 (SEQ ID NO: 78)

```
v.1          1 TCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTCTTTTATT     50
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1 TCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTCTTTTATT     50

v.1         51 CATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGTGCTCTCA    100
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4         51 CATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGTGCTCTCA    100

v.1        101 TCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAGATGGCAC    150
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        101 TCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAGATGGCAC    150

v.1        151 AATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCTGAAATAA    200
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        151 AATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCTGAAATAA    200

v.1        201 GTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAACGGCTTG    250
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        201 GTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAACGGCTTG    250

v.1        251 ACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT    300
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        251 ACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT    300

v.1        301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATG    350
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATG    350

v.1        351 GCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATT    400
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        351 GCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATT    400

v.1        401 CTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCCTGCAAGC    450
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        401 CTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCCTGCAAGC    450

v.1        451 AGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGCAAGCTCA    500
               ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4        451 AGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGCAAGCTCA    500

v.1        501 ACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAGTCTTCCT    550
```


v.4 501 ACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAGTCTTCCT 550
v.1 551 CCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTCGTGGAAA 600
v.4 551 CCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTCGTGGAAA 600
v.1 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAA 650
v.4 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAA 650
v.1 651 TATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTA 700
v.4 651 TATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTA 700
v.1 701 TTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTATAATTGG 750
v.4 701 TTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTATAATTGG 750
v.1 751 TGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATACTCAGTA 800
v.4 751 TGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATACTCAGTA 800
v.1 801 GACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGAAGAACAT 850
v.4 801 GACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGAAGAACAT 850
v.1 851 GAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA 900
v.4 851 GAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA 900
v.1 901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAG 950
v.4 901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAG 950
v.1 951 CACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGA 1000
v.4 951 CACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGA 1000
v.1 1001 CCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCATCAGGACT 1050
v.4 1001 CCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCATCAGGACT 1050
v.1 1051 TCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGATCTGAGAC 1100
v.4 1051 TCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGATCTGAGAC 1100
v.1 1101 CTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATATTATTCAC 1150
v.4 1101 CTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATATTATTCAC 1150
v.1 1151 AGTTTAATGAAGTCTGATCTAGTGGAATATTTCACTTTGGAAATGCTTCA 1200
v.4 1151 AGTTTAATGAAGTC---------- 1164
v.1 1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACC 1250
v.4 1165 ---------- 1164
v.1 1251 TAACGAGATTACAAAAACTCTATCTAAATGGTAACCACCTGACCAAATTA 1300
v.4 1165 ---------- 1164
v.1 1301 AGTAAAGGCATGTTCCTTGGTCTCCATAATCTTGAATACTTATATCTTGA 1350
v.4 1165 ---------- 1164
v.1 1351 ATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCAATGCCTA 1400
v.4 1165 ---------- 1164

```
v.1          1401 AACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTTACCACCA   1450
v.4          1165 ----------------------------------------------------   1164
v.1          1451 CATATTTTTTCAGGGGTTCCTCTAACTAAGGTAAATCTTAAAACAAACCA   1500
v.4          1165 ----------------------------------------------------   1164
v.1          1501 GTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGATTTACTAA   1550
v.4          1165 ----------------------------------------------------   1164
v.1          1551 CCCAGATTGACCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTT   1600
v.4          1165 ----------------------------------------------------   1164
v.1          1601 GGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGACAGATGA   1650
v.4          1165 ----------------------------------------------------   1164
v.1          1651 CATCCTCTGCACTTCCCCCGGGCATCTCGACAAAAAGGAATTGAAAGCCC   1700
v.4          1165 ----------------------------------------------------   1164
v.1          1701 TAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATCCATGCCA   1750
v.4          1165 ----------------------------------------------------   1164
v.1          1751 ACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC   1800
v.4          1165 ----------------------------------------------------   1164
v.1          1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTC   1850
v.4          1165 ----------------------------------------------------   1164
v.1          1851 TAATATTGGGACTTCTGATTATGTTCATCACTATTGTTTTCTGTGCTGCA   1900
v.4          1165 ----------------------------------------------------   1164
v.1          1901 GGGATAGTGGTTCTTGTTCTTCACCGCAGGAGAAGATACAAAAAGAAACA   1950
v.4          1165 ----------------------------------------------------   1164
v.1          1951 AGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAGTACAGCA   2000
v.4          1165 ----------------------------------------------------   1164
v.1          2001 TGTATGGCCATAAAACCACTCATCACACTACTGAAAGACCCTCTGCCTCA   2050
v.4          1165 ----------------------------------------------------   1164
v.1          2051 CTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATAGAAGTCC   2100
                                                                 |
v.4          1165 ----------------------------------------------C   1165
v.1          2101 ATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAATGAGAAAG   2150
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1166 ATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAATGAGAAAG   1215
v.1          2151 AAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACAGGAAAAT   2200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1216 AAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACAGGAAAAT   1265
v.1          2201 CATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGAACCAATC   2250
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1266 CATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGAACCAATC   1315
v.1          2251 AACAGAATTTTTATCCTTCCAAGATGCCAGCTCATTGTACAGAAACATTT   2300
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1316 AACAGAATTTTTATCCTTCCAAGATGCCAGCTCATTGTACAGAAACATTT   1365

v.1          2301 TAGAAAAAGAAAGGGAACTTCAGCAACTGGGAATCACAGAATACCTAAGG   2350
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1366 TAGAAAAAGAAAGGGAACTTCAGCAACTGGGAATCACAGAATACCTAAGG   1415

v.1          2351 AAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGC   2400
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1416 AAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGC   1465

v.1          2401 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGA   2450
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1466 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGA   1515

v.1          2451 AGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAAAGCTAAT   2500
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1516 AGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAAAGCTAAT   1565

v.1          2501 TTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATAGAT   2550
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4          1566 TTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATAGAT   1615

v.1          2551 GGAGA   2555
                  |||||
v.4          1616 GGAGA   1620
```

Table LIII(b). Peptide sequences of protein coded by 158P1D7 v.4 (SEQ ID NO: 79)

```
MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLCNCEEK DGTMLINCEA KGIKMVSEIS      60
VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIHLGFNNI ADIEIGAFNG LGLLKQLHIN     120
HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF     180
RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS     240
IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT     300
TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP     360
PPQNPRKLIL AGNIIHSLMK SILWSKASGR GRREE                                395
```

Table LIV(b). Amino acid sequence alignment of 158P1D7 v.1 (SEQ ID NO: 80) and 158P1D7 v.4 (SEQ ID NO: 81)

```
v.1             1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4             1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50

v.1            51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4            51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100

v.1           101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4           101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150

v.1           151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4           151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200

v.1           201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4           201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250

v.1           251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4           251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300

v.1           301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.4           301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350

v.1           351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400
                  ||||||||||||||||||||||||||||||.|
v.4           351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSIL----------------    383
```

```
v.1          401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450

v.4          384 ----------------------------------------------------    383

v.1          451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500

v.4          384 ----------------------------------------------------    383

v.1          501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550
                                 |                ||
v.4          384 ------------------W----------------SK-------------    386

v.1          551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600

v.4          387 ----------------------------------------------------    386

v.1          601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650
                                      |:|       |.||
v.4          387 ----------------------ASG-------RGRR-------------    393

v.1          651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHL    700

v.4          394 ----------------------------------------------------    393

v.1          701 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    750

v.4          394 ----------------------------------------------------    393

v.1          751 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    800

v.4          394 ----------------------------------------------------    393

v.1          801 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    841
                                                        :.
v.4          394 ----------------------------------------EE    395
```

Table LI(c). Nucleotide sequence of transcript variant 158P1D7 v.5 (SEQ ID NO: 82)

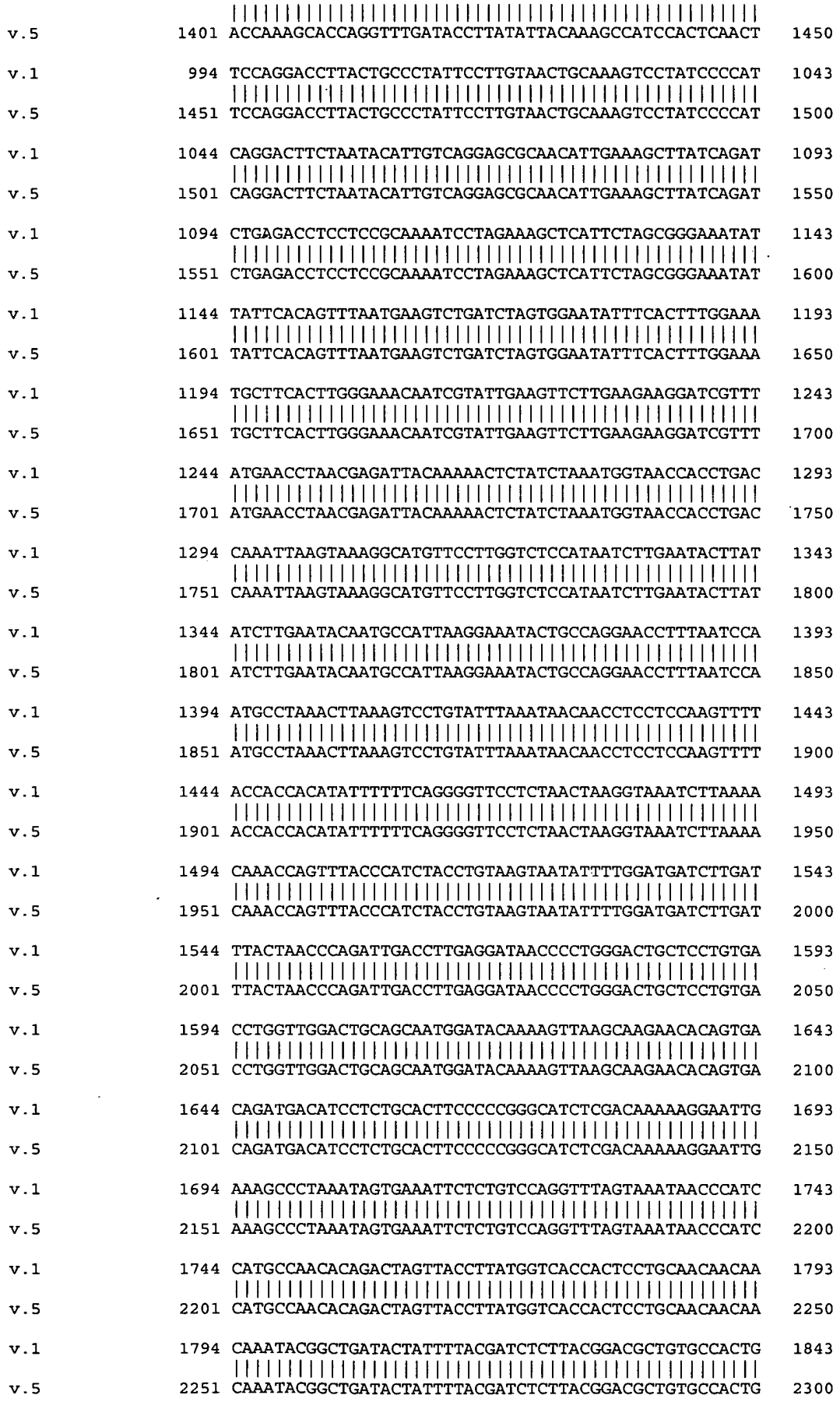
```
gcgtcgacaa caagaaatac tagaaaagga ggaaggagaa cattgctgca gcttggatct      60
acaacctaag aaagcaagag tgatcaatct cagctctgtt aaacatcttg tttacttact     120
gcattcagca gcttgcaaat ggttaactat atgcaaaaaa gtcagcatag ctgtgaagta     180
tgccgtgaat tttaattgag ggaaaaagga caattgcttc aggatgctct agtatgcact     240
ctgcttgaaa tattttcaat gaaatgctca gtattctatc tttgaccaga ggttttaact     300
ttatgaagct atgggacttg acaaaaagtg atatttgaga agaaagtacg cagtggttgg     360
tgtttttctt tttttaataa aggaattgaa ttactttgaa cacctcttcc agctgtgcat     420
tacagataac gtcaggaaga gtctctgctt tacagaatcg gatttcatca catgacaaca     480
tgaagctgtg gattcatctc ttttattcat ctctccttgc ctgtatatct ttacactccc     540
aaactccagt gctctcatcc agaggctctt gtgattctct ttgcaattgt gaggaaaaag     600
atggcacaat gctaataaat tgtgaagcaa aaggtatcaa gatggtatct gaaataagtg     660
tgccaccatc acgacctttc caactaagct tattaaataa cggcttgacg atgcttcaca     720
caaatgactt ttctgggctt accaatgcta tttcaataca ccttggattt aacaatattg     780
cagatattga gataggtgca tttaatggcc ttggcctcct gaaacaactt catatcaatc     840
acaattcttt agaaattctt aaagaggata ctttccatgg actggaaaac ctggaattcc     900
tgcaagcaga taacaatttt atcacagtga ttgaaccaag tgcctttagc aagctcaaca     960
gactcaaagt gttaatttta aatgacaatg ctattgagag tcttcctcca aacatcttcc    1020
gatttgttcc tttaacccat ctagatcttc gtggaaatca attacaaaca ttgccttatg    1080
ttggttttct cgaacacatt ggccgaatat tggatcttca gttggaggac aacaaatggg    1140
cctgcaattg tgacttattg cagttaaaaa cttggttgga gaacatgcct ccacagtcta    1200
taattggtga tgttgtctgc aacagccctc catttttaa aggaagtata ctcagtagac     1260
taaagaagga atctatttgc cctactccac cagtgtatga agaacatgag gatccttcag    1320
gatcattaca tctggcagca acatcttcaa taaatgatag tcgcatgtca actaagacca    1380
cgtccattct aaaactaccc accaaagcac caggtttgat accttatatt acaaagccat    1440
ccactcaact tccaggacct tactgcccta ttccttgtaa ctgcaaagtc ctatccccat    1500
caggacttct aatacattgt caggagcgca acattgaaag cttatcagat ctgagacctc    1560
ctccgcaaaa tcctagaaag ctcattctag cgggaaatat tattcacagt ttaatgaagt    1620
ctgatctagt ggaatatttc actttggaaa tgcttcactt gggaaacaat cgtattgaag    1680
ttcttgaaga aggatcgttt atgaacctaa cgagattaca aaaactctat ctaaatggta    1740
accacctgac caaattaagt aaaggcatgt tccttggtct ccataatctt gaatacttat    1800
atcttgaata caatgccatt aaggaaatac tgccaggaac ctttaatcca atgcctaaac    1860
```

```
ttaaagtcct gtatttaaat aacaacctcc tccaagtttt accaccacat atttttcag    1920
gggttcctct aactaaggta aatcttaaaa caaaccagtt tacccatcta cctgtaagta    1980
atattttgga tgatcttgat ttactaaccc agattgacct tgaggataac ccctgggact    2040
gctcctgtga cctggttgga ctgcagcaat ggatacaaaa gttaagcaag aacacagtga    2100
cagatgacat cctctgcact tcccccgggc atctcgacaa aaaggaattg aaagccctaa    2160
atagtgaaat tctctgtcca ggtttagtaa ataacccatc catgccaaca cagactagtt    2220
accttatggt caccactcct gcaacaacaa caaatacggc tgatactatt ttacgatctc    2280
ttacggacgc tgtgccactg tctgttctaa tattgggact tctgattatg ttcatcacta    2340
ttgttttctg tgctgcaggg atagtggttc ttgttcttca ccgcaggaga agatacaaaa    2400
agaaacaagt agatgagcaa atgagagaca acagtcctgt gcatcttcag tacagcatgt    2460
atggccataa aaccactcat cacactactg aaagaccctc tgcctcactc tatgaacagc    2520
acatggtgag ccccatggtt catgtctata gaagtccatc ctttggtcca aagcatctgg    2580
aagaggaaga agagaggaat gagaaagaag gaagtgatgc aaaacatctc caaagaagtc    2640
ttttggaaca ggaaaatcat tcaccactca cagggtcaaa tatgaaatac aaaaccacga    2700
accaatcaac agaattttta tccttccaag atgccagctc attgtacaga aacattttag    2760
aaaaagaaag ggaacttcag caactgggaa tcacagaata cctaaggaaa aacattgctc    2820
agctccagcc tgatatggag gcacattatc ctggagccca cgaagagctg aagttaatgg    2880
aaacattaat gtactcacgt ccaaggaagg tattagtgga acagacaaaa aatgagtatt    2940
ttgaacttaa agctaattta catgctgaac ctgactattt agaagtcctg gagcagcaaa    3000
catagatgga gagttgaggg ctttcgccag aaatgctgtg attctgttat taagtccata    3060
ccttgtaaat aagtgcctta cgtgagtgtg tcatcaatca gaacctaagc acagagtaaa    3120
ctatggggaa aaaaaaagaa gacgaaacag aaactcaggg atcactggga gaagccatgg    3180
cataatcttc aggcaattta gtctgtccca aataaacata catccttggc atgtaaatca    3240
tcaagggtaa tagtaatatt catatacctg aaacgtgtct cataggagtc ctctctgcac    3300
```

Table LII(c). Nucleotide sequence alignment of 158P1D7 v.1 (SEQ ID NO: 83) and 158P1D7 v.5 (SEQ ID NO: 84)

```
v.1                1 --------------------------------------------------      0

v.5                1 GCGTCGACAACAAGAAATACTAGAAAAGGAGGAAGGAGAACATTGCTGCA     50

v.1                1 --------------------------------------------------      0

v.5               51 GCTTGGATCTACAACCTAAGAAAGCAAGAGTGATCAATCTCAGCTCTGTT    100

v.1                1 --------------------------------------------------      0

v.5              101 AAACATCTTGTTTACTTACTGCATTCAGCAGCTTGCAAATGGTTAACTAT    150

v.1                1 --------------------------------------------------      0

v.5              151 ATGCAAAAAAGTCAGCATAGCTGTGAAGTATGCCGTGAATTTTAATTGAG    200

v.1                1 --------------------------------------------------      0

v.5              201 GGAAAAAGGACAATTGCTTCAGGATGCTCTAGTATGCACTCTGCTTGAAA    250

v.1                1 --------------------------------------------------      0

v.5              251 TATTTTCAATGAAATGCTCAGTATTCTATCTTTGACCAGAGGTTTTAACT    300

v.1                1 --------------------------------------------------      0

v.5              301 TTATGAAGCTATGGGACTTGACAAAAAGTGATATTTGAGAAGAAAGTACG    350

v.1                1 --------------------------------------------------      0

v.5              351 CAGTGGTTGGTGTTTTCTTTTTTTTAATAAAGGAATTGAATTACTTTGAA    400

v.1                1 --------------------------------------------------      0

v.5              401 CACCTCTTCCAGCTGTGCATTACAGATAACGTCAGGAAGAGTCTCTGCTT    450

v.1                1 -------TCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTC     43
                             |||||||||||||||||||||||||||||||||||||||||
v.5              451 TACAGAATCGGATTTCATCACATGACAACATGAAGCTGTGGATTCATCTC    500

v.1               44 TTTTATTCATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGT     93
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.5              501 TTTTATTCATCTCTCCTTGCCTGTATATCTTTACACTCCCAAACTCCAGT    550
```

```
v.1          94 GCTCTCATCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAG  143
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         551 GCTCTCATCCAGAGGCTCTTGTGATTCTCTTTGCAATTGTGAGGAAAAAG  600

v.1         144 ATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCT  193
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         601 ATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCT  650

v.1         194 GAAATAAGTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAA  243
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         651 GAAATAAGTGTGCCACCATCACGACCTTTCCAACTAAGCTTATTAAATAA  700

v.1         244 CGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTA  293
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         701 CGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTA  750

v.1         294 TTTCAATACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCA  343
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         751 TTTCAATACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCA  800

v.1         344 TTTAATGGCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTT  393
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         801 TTTAATGGCCTTGGCCTCCTGAAACAACTTCATATCAATCACAATTCTTT  850

v.1         394 AGAAATTCTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCC  443
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         851 AGAAATTCTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCC  900

v.1         444 TGCAAGCAGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGC  493
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         901 TGCAAGCAGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGC  950

v.1         494 AAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAG  543
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5         951 AAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAG 1000

v.1         544 TCTTCCTCCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTC  593
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1001 TCTTCCTCCAAACATCTTCCGATTTGTTCCTTTAACCCATCTAGATCTTC 1050

v.1         594 GTGGAAATCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATT  643
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1051 GTGGAAATCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATT 1100

v.1         644 GGCCGAATATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTG  693
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1101 GGCCGAATATTGGATCTTCAGTTGGAGGACAACAAATGGGCCTGCAATTG 1150

v.1         694 TGACTTATTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTA  743
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1151 TGACTTATTGCAGTTAAAAACTTGGTTGGAGAACATGCCTCCACAGTCTA 1200

v.1         744 TAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATA  793
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1201 TAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATA 1250

v.1         794 CTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGA  843
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1251 CTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGA 1300

v.1         844 AGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAA  893
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1301 AGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAA 1350

v.1         894 TAAATGATAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCC  943
                ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5        1351 TAAATGATAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCC 1400

v.1         944 ACCAAAGCACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACT  993
```


v.5 1401 ACCAAAGCACCAGGTTTGATACCTTATATTACAAAGCCATCCACTCAACT 1450
v.1 994 TCCAGGACCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCAT 1043
v.5 1451 TCCAGGACCTTACTGCCCTATTCCTTGTAACTGCAAAGTCCTATCCCCAT 1500
v.1 1044 CAGGACTTCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGAT 1093
v.5 1501 CAGGACTTCTAATACATTGTCAGGAGCGCAACATTGAAAGCTTATCAGAT 1550
v.1 1094 CTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATAT 1143
v.5 1551 CTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAATAT 1600
v.1 1144 TATTCACAGTTTAATGAAGTCTGATCTAGTGGAATATTTCACTTTGGAAA 1193
v.5 1601 TATTCACAGTTTAATGAAGTCTGATCTAGTGGAATATTTCACTTTGGAAA 1650
v.1 1194 TGCTTCACTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTT 1243
v.5 1651 TGCTTCACTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTT 1700
v.1 1244 ATGAACCTAACGAGATTACAAAAACTCTATCTAAATGGTAACCACCTGAC 1293
v.5 1701 ATGAACCTAACGAGATTACAAAAACTCTATCTAAATGGTAACCACCTGAC 1750
v.1 1294 CAAATTAAGTAAAGGCATGTTCCTTGGTCTCCATAATCTTGAATACTTAT 1343
v.5 1751 CAAATTAAGTAAAGGCATGTTCCTTGGTCTCCATAATCTTGAATACTTAT 1800
v.1 1344 ATCTTGAATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCA 1393
v.5 1801 ATCTTGAATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCA 1850
v.1 1394 ATGCCTAAACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTT 1443
v.5 1851 ATGCCTAAACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTT 1900
v.1 1444 ACCACCACATATTTTTTCAGGGGTTCCTCTAACTAAGGTAAATCTTAAAA 1493
v.5 1901 ACCACCACATATTTTTTCAGGGGTTCCTCTAACTAAGGTAAATCTTAAAA 1950
v.1 1494 CAAACCAGTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGAT 1543
v.5 1951 CAAACCAGTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGAT 2000
v.1 1544 TTACTAACCCAGATTGACCTTGAGGATAACCCCTGGGACTGCTCCTGTGA 1593
v.5 2001 TTACTAACCCAGATTGACCTTGAGGATAACCCCTGGGACTGCTCCTGTGA 2050
v.1 1594 CCTGGTTGGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGA 1643
v.5 2051 CCTGGTTGGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGA 2100
v.1 1644 CAGATGACATCCTCTGCACTTCCCCCGGGCATCTCGACAAAAAGGAATTG 1693
v.5 2101 CAGATGACATCCTCTGCACTTCCCCCGGGCATCTCGACAAAAAGGAATTG 2150
v.1 1694 AAAGCCCTAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATC 1743
v.5 2151 AAAGCCCTAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATC 2200
v.1 1744 CATGCCAACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAA 1793
v.5 2201 CATGCCAACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAA 2250
v.1 1794 CAAATACGGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTG 1843
v.5 2251 CAAATACGGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTG 2300

```
v.1          1844 TCTGTTCTAATATTGGGACTTCTGATTATGTTCATCACTATTGTTTTCTG 1893
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2301 TCTGTTCTAATATTGGGACTTCTGATTATGTTCATCACTATTGTTTTCTG 2350

v.1          1894 TGCTGCAGGGATAGTGGTTCTTGTTCTTCACCGCAGGAGAAGATACAAAA 1943
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2351 TGCTGCAGGGATAGTGGTTCTTGTTCTTCACCGCAGGAGAAGATACAAAA 2400

v.1          1944 AGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAG 1993
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2401 AGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAG 2450

v.1          1994 TACAGCATGTATGGCCATAAAACCACTCATCACACTACTGAAAGACCCTC 2043
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2451 TACAGCATGTATGGCCATAAAACCACTCATCACACTACTGAAAGACCCTC 2500

v.1          2044 TGCCTCACTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATA 2093
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2501 TGCCTCACTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATA 2550

v.1          2094 GAAGTCCATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAAT 2143
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2551 GAAGTCCATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAAT 2600

v.1          2144 GAGAAAGAAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACA 2193
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2601 GAGAAAGAAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACA 2650

v.1          2194 GGAAAATCATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGA 2243
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2651 GGAAAATCATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGA 2700

v.1          2244 ACCAATCAACAGAATTTTTATCCTTCCAAGATGCCAGCTCATTGTACAGA 2293
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2701 ACCAATCAACAGAATTTTTATCCTTCCAAGATGCCAGCTCATTGTACAGA 2750

v.1          2294 AACATTTTAGAAAAAGAAAGGGAACTTCAGCAACTGGGAATCACAGAATA 2343
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2751 AACATTTTAGAAAAAGAAAGGGAACTTCAGCAACTGGGAATCACAGAATA 2800

v.1          2344 CCTAAGGAAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATC 2393
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2801 CCTAAGGAAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATC 2850

v.1          2394 CTGGAGCCCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGT 2443
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2851 CTGGAGCCCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGT 2900

v.1          2444 CCAAGGAAGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAA 2493
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2901 CCAAGGAAGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACTTAA 2950

v.1          2494 AGCTAATTTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAA 2543
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          2951 AGCTAATTTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAA 3000

v.1          2544 CATAGATGGAGA---------------------------------------- 2555
                  ||||||||||||
v.5          3001 CATAGATGGAGAGTTGAGGGCTTTCGCCAGAAATGCTGTGATTCTGTTAT 3050

v.1          2556 ---------------------------------------------------- 2555

v.5          3051 TAAGTCCATACCTTGTAAATAAGTGCCTTACGTGAGTGTGTCATCAATCA 3100

v.1          2556 ---------------------------------------------------- 2555

v.5          3101 GAACCTAAGCACAGAGTAAACTATGGGGAAAAAAAAAAGAAGACGAAACAG 3150

v.1          2556 ---------------------------------------------------- 2555
```

```
v.5          3151 AAACTCAGGGATCACTGGGAGAAGCCATGGCATAATCTTCAGGCAATTTA   3200

v.1          2556 ------------------------------------------------   2555

v.5          3201 GTCTGTCCCAAATAAACATACATCCTTGGCATGTAAATCATCAAGGGTAA   3250

v.1          2556 ------------------------------------------------   2555

v.5          3251 TAGTAATATTCATATACCTGAAACGTGTCTCATAGGAGTCCTCTCTGCAC   330
```

Table LIII(c). Peptide sequences of protein coded by 158P1D7 v.5 (SEQ ID NO: 85)

```
MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLCNCEEK DGTMLINCEA KGIKMVSEIS      60
VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIHLGFNNI ADIEIGAFNG LGLLKQLHIN     120
HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF     180
RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS     240
IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT     300
TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP     360
PPQNPRKLIL AGNIIHSLMK SDLVEYFTLE MLHLGNNRIE VLEEGSFMNL TRLQKLYLNG     420
NHLTKLSKGM FLGLHNLEYL YLEYNAIKEI LPGTFNPMPK LKVLYLNNNL LQVLPPHIFS     480
GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIQKLSKNTV     540
TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS     600
LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM     660
YGHKTTHHTT ERPSASLYEQ HMVSPMVHVY RSPSFGPKHL EEEEERNEKE GSDAKHLQRS     720
LLEQENHSPL TGSNMKYKTT NQSTEFLSFQ DASSLYRNIL EKERELQQLG ITEYLRKNIA     780
QLQPDMEAHY PGAHEELKLM ETLMYSRPRK VLVEQTKNEY FELKANLHAE PDYLEVLEQQ     840
T                                                                     841
```

Table LIV(c). Amino acid sequence alignment of 158P1D7 v.1 (SEQ ID NO: 86) and 158P1D7 v.5 (SEQ ID NO: 87)

```
v.1             1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5             1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50

v.1            51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5            51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100

v.1           101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150

v.1           151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200

v.1           201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250

v.1           251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300

v.1           301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350

v.1           351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400

v.1           401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450

v.1           451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
v.5           451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500
```

```
v.1          501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550

v.1          551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600

v.1          601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650

v.1          651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHL    700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHL    700

v.1          701 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    750
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          701 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    750

v.1          751 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    800
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
v.5          751 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    800

v.1          801 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    841
                 |||||||||||||||||||||||||||||||||||||||||
v.5          801 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    841
```

Table LI(d). Nucleotide sequence of transcript variant 158P1D7 v.6 (SEQ ID NO: 88)

```
tcggatttca tcacatgaca acatgaagct gtggattcat ctcttttatt catctctcct     60
tgcctgtata tctttacact cccaaactcc agtgctctca tccagaggct cttgtgattc    120
tctttgcaat tgtgaggaaa aagatggcac aatgctaata aattgtgaag caaaaggtat    180
caagatggta tctgaaataa gtgtgccacc atcacgacct ttccaactaa gcttattaaa    240
taacggcttg acgatgcttc acacaaatga cttttctggg cttaccaatg ctatttcaat    300
acaccttgga tttaacaata ttgcagatat tgagataggt gcatttaatg gccttggcct    360
cctgaaacaa cttcatatca atcacaattc tttagaaatt cttaaagagg atactttcca    420
tggactggaa aacctggaat tcctgcaagc agataacaat tttatcacag tgattgaacc    480
aagtgccttt agcaagctca acagactcaa agtgttaatt ttaaatgaca atgctattga    540
gagtcttcct ccaaacatct tccgatttgt tcctttaacc catctagatc ttcgtggaaa    600
tcaattacaa acattgcctt atgttggttt tctcgaacac attggccgaa tattggatct    660
tcagttggag gacaacaaat gggcctgcaa ttgtgactta ttgcagttaa aaacttggtt    720
ggagaacatg cctccacagt ctataattgg tgatgttgtc tgcaacagcc ctccattttt    780
taaaggaagt atactcagta gactaaagaa ggaatctatt tgccctactc caccagtgta    840
tgaagaacat gaggatcctt caggatcatt acatctggca gcaacatctt caataaatga    900
tagtcgcatg tcaactaaga ccacgtccat tctaaaacta cccaccaaag caccaggttt    960
gataccttat attacaaagc catccactca acttccagga ccttactgcc ctattccttg   1020
taactgcaaa gtcctatccc catcaggact tctaatacat tgtcaggagc gcaacattga   1080
aagcttatca gatctgagac ctcctccgca aaatcctaga aagctcattc tagcgggaaa   1140
tattattcac agtttaatga atccatcctt tggtccaaag catctggaag aggaagaaga   1200
gaggaatgag aaagaaggaa gtgatgcaaa acatctccaa agaagtcttt tggaacagga   1260
aaatcattca ccactcacag ggtcaaatat gaaatacaaa accacgaacc aatcaacaga   1320
atttttatcc ttccaagatg ccagctcatt gtacagaaac attttagaaa aagaaaggga   1380
acttcagcaa ctgggaatca cagaatacct aaggaaaaac attgctcagc tccagcctga   1440
tatggaggca cattatcctg gagcccacga agagctgaag ttaatggaaa cattaatgta   1500
ctcacgtcca aggaaggtat tagtggaaca gacaaaaaat gagtattttg aacttaaagc   1560
taatttacat gctgaacctg actatttaga agtcctggag cagcaaacat agatggaga    1619
```

Table LII(d). Nucleotide sequence alignment of 158P1D7 v.1 (SEQ ID NO: 89) and 158P1D7 v.6 (SEQ ID NO: 90)

```
V.1  : 1    tcggatttcatcacatgacaacatgaagctgtggattcatctcttttattcatctctcct 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1    tcggatttcatcacatgacaacatgaagctgtggattcatctcttttattcatctctcct 60


V.1  : 61   tgcctgtatatctttacactcccaaactccagtgctctcatccagaggctcttgtgattc 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 61   tgcctgtatatctttacactcccaaactccagtgctctcatccagaggctcttgtgattc 120
```

```
V.1  : 121  tctttgcaattgtgaggaaaaagatggcacaatgctaataaattgtgaagcaaaaggtat 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 121  tctttgcaattgtgaggaaaaagatggcacaatgctaataaattgtgaagcaaaaggtat 180

V.1  : 181  caagatggtatctgaaataagtgtgccaccatcacgacctttccaactaagcttattaaa 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 181  caagatggtatctgaaataagtgtgccaccatcacgacctttccaactaagcttattaaa 240

V.1  : 241  taacggcttgacgatgcttcacacaaatgacttttctgggcttaccaatgctatttcaat 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 241  taacggcttgacgatgcttcacacaaatgacttttctgggcttaccaatgctatttcaat 300

V.1  : 301  acaccttggatttaacaatattgcagatattgagataggtgcatttaatggccttggcct 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 301  acaccttggatttaacaatattgcagatattgagataggtgcatttaatggccttggcct 360

V.1  : 361  cctgaaacaacttcatatcaatcacaattctttagaaattcttaaagaggatactttcca 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 361  cctgaaacaacttcatatcaatcacaattctttagaaattcttaaagaggatactttcca 420

V.1  : 421  tggactggaaaacctggaattcctgcaagcagataacaattttatcacagtgattgaacc 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 421  tggactggaaaacctggaattcctgcaagcagataacaattttatcacagtgattgaacc 480

V.1  : 481  aagtgcctttagcaagctcaacagactcaaagtgttaattttaaatgacaatgctattga 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 481  aagtgcctttagcaagctcaacagactcaaagtgttaattttaaatgacaatgctattga 540

V.1  : 541  gagtcttcctccaaacatcttccgatttgttcctttaacccatctagatcttcgtggaaa 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 541  gagtcttcctccaaacatcttccgatttgttcctttaacccatctagatcttcgtggaaa 600

V.1  : 601  tcaattacaaacattgccttatgttggttttctcgaacacattggccgaatattggatct 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 601  tcaattacaaacattgccttatgttggttttctcgaacacattggccgaatattggatct 660

V.1  : 661  tcagttggaggacaacaaatgggcctgcaattgtgacttattgcagttaaaaacttggtt 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 661  tcagttggaggacaacaaatgggcctgcaattgtgacttattgcagttaaaaacttggtt 720

V.1  : 721  ggagaacatgcctccacagtctataattggtgatgttgtctgcaacagccctccatttt 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 721  ggagaacatgcctccacagtctataattggtgatgttgtctgcaacagccctccatttt 780

V.1  : 781  taaaggaagtatactcagtagactaaagaaggaatctatttgccctactccaccagtgta 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 781  taaaggaagtatactcagtagactaaagaaggaatctatttgccctactccaccagtgta 840

V.1  : 841  tgaagaacatgaggatccttcaggatcattacatctggcagcaacatcttcaataaatga 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 841  tgaagaacatgaggatccttcaggatcattacatctggcagcaacatcttcaataaatga 900

V.1  : 901  tagtcgcatgtcaactaagaccacgtccattctaaaactacccaccaaagcaccaggttt 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 901  tagtcgcatgtcaactaagaccacgtccattctaaaactacccaccaaagcaccaggttt 960
```

```
V.1  : 961  gataccttatattacaaagccatccactcaacttccaggaccttactgccctattccttg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 961  gataccttatattacaaagccatccactcaacttccaggaccttactgccctattccttg 1020

V.1  : 1021 taactgcaaagtcctatccccatcaggacttctaatacattgtcaggagcgcaacattga 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1021 taactgcaaagtcctatccccatcaggacttctaatacattgtcaggagcgcaacattga 1080

V.1  : 1081 aagcttatcagatctgagacctcctccgcaaaatcctagaaagctcattctagcgggaaa 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1081 aagcttatcagatctgagacctcctccgcaaaatcctagaaagctcattctagcgggaaa 1140

V.1  : 1141 tattattcacagtttaatgaa 1161
           |||||||||||||||||||||
V.6  : 1141 tattattcacagtttaatgaa 1161

V.1  : 2098 tccatcctttggtccaaagcatctggaagaggaagaagagaggaatgagaaagaaggaag 2157
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1162 tccatcctttggtccaaagcatctggaagaggaagaagagaggaatgagaaagaaggaag 1221

V.1  : 2158 tgatgcaaaacatctccaaagaagtcttttggaacaggaaaatcattcaccactcacagg 2217
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1222 tgatgcaaaacatctccaaagaagtcttttggaacaggaaaatcattcaccactcacagg 1281

V.1  : 2218 gtcaaatatgaaatacaaaaccacgaaccaatcaacagaatttttatccttccaagatgc 2277
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1282 gtcaaatatgaaatacaaaaccacgaaccaatcaacagaatttttatccttccaagatgc 1341

V.1  : 2278 cagctcattgtacagaaacattttagaaaaagaaagggaacttcagcaactgggaatcac 2337
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1342 cagctcattgtacagaaacattttagaaaaagaaagggaacttcagcaactgggaatcac 1401

V.1  : 2338 agaatacctaaggaaaaacattgctcagctccagcctgatatggaggcacattatcctgg 2397
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1402 agaatacctaaggaaaaacattgctcagctccagcctgatatggaggcacattatcctgg 1461

V.1  : 2398 agcccacgaagagctgaagttaatggaaacattaatgtactcacgtccaaggaaggtatt 2457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1462 agcccacgaagagctgaagttaatggaaacattaatgtactcacgtccaaggaaggtatt 1521

V.1  : 2458 agtggaacagacaaaaaatgagtattttgaacttaaagctaatttacatgctgaacctga 2517
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
V.6  : 1522 agtggaacagacaaaaaatgagtattttgaacttaaagctaatttacatgctgaacctga 1581

V.1  : 2518 ctatttagaagtcctggagcagcaaacatagatggaga 2555
           |||||||||||||||||||||||||||||||||||||
V.6  : 1582 ctatttagaagtcctggagcagcaaacatagatggaga 1619
```

Table LIII(d). Peptide sequences of protein coded by 158P1D7 v.6 (SEQ ID NO: 91)

```
MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLCNCEEK DGTMLINCEA KGIKMVSEIS      60
VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIHLGFNNI ADIEIGAFNG LGLLKQLHIN     120
HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF     180
RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS     240
IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT     300
TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP     360
PPQNPRKLIL AGNIIHSLMN PSFGPKHLEE EEERNEKEGS DAKHLQRSLL EQENHSPLTG     420
```

```
SNMKYKTTNQ STEFLSFQDA SSLYRNILEK ERELQQLGIT EYLRKNIAQL QPDMEAHYPG     480
AHEELKLMET LMYSRPRKVL VEQTKNEYFE LKANLHAEPD YLEVLEQQT                529
```

Table LIV(d). Amino acid sequence alignment of 158P1D7 v.1 (SEQ ID NO: 92) and 158P1D7 v.6 (SEQ ID NO: 93)

```
v.1                1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6                1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLCNCEEKDGTMLINCEA     50

v.1               51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6               51 KGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIHLGFNNI    100

v.1              101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              101 ADIEIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITV    150

v.1              151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              151 IEPSAFSKLNRLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPY    200

v.1              201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              201 VGFLEHIGRILDLQLEDNKWACNCDLLQLKTWLENMPPQSIIGDVVCNSP    250

v.1              251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              251 PFFKGSILSRLKKESICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKT    300

v.1              301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              301 TSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGLLIHCQER    350

v.1              351 NIESLSDLRPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIE    400
                     |||||||||||||||||||||||||||
v.6              351 NIESLSDLRPPPQNPRKLILAGNIIHSLM----------------------    379

v.1              401 VLEEGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEI    450

v.6              380 --------------------------------------------------    379

v.1              451 LPGTFNPMPKLKVLYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVS    500

v.6              380 --------------------------------------------------    379

v.1              501 NILDDLDLLTQIDLEDNPWDCSCDLVGLQQWIQKLSKNTVTDDILCTSPG    550

v.6              380 --------------------------------------------------    379

v.1              551 HLDKKELKALNSEILCPGLVNNPSMPTQTSYLMVTTPATTTNTADTILRS    600

v.6              380 --------------------------------------------------    379

v.1              601 LTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLHRRRRYKKKQVDEQMRD    650

v.6              380 --------------------------------------------------    379

v.1              651 NSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHL    700
                                                        :||||||||
v.6              380 ------------------------------------------NPSFGPKHL    388

v.1              701 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    750
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              389 EEEEERNEKEGSDAKHLQRSLLEQENHSPLTGSNMKYKTTNQSTEFLSFQ    438

v.1              751 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    800
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
v.6              439 DASSLYRNILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLM    488

v.1              801 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT    841
                     |||||||||||||||||||||||||||||||||||||||||
```

v.6 489 ETLMYSRPRKVLVEQTKNEYFELKANLHAEPDYLEVLEQQT 529

Table LV: Search peptides

158P1D7,variant 1: 9-mers 10-mers and 15-mers (SEQ ID NO: 94)

MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLCNCEEK DGTMLINCEA KGIKMVSEIS
VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIHLGFNNI ADIEIGAFNG LGLLKQLHIN
HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF
RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS
IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT
TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP
PPQNPRKLIL AGNIIHSLMK SDLVEYFTLE MLHLGNNRIE VLEEGSFMNL TRLQKLYLNG
NHLTKLSKGM FLGLHNLEYL YLEYNAIKEI LPGTFNPMPK LKVLYLNNNL LQVLPPHIFS
GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIQKLSKNTV
TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS
LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM
YGHKTTHHTT ERPSASLYEQ HMVSPMVHVY RSPSFGPKHL EEEEERNEKE GSDAKHLQRS
LLEQENHSPL TGSNMKYKTT NQSTEFLSFQ DASSLYRNIL EKERELQQLG ITEYLRKNIA
QLQPDMEAHY PGAHEELKLM ETLMYSRPRK VLVEQTKNEY FELKANLHAE PDYLEVLEQQ
T

158P1D7 Variant 3:
9-mers
ASLYEQHMGAHEELKL (SEQ ID NO: 95)start position 675
10-mers
SASLYEQHMGAHEELKLM (SEQ ID NO: 96) start position 674
15-mers
TTERPSASLYEQHMGAHEELKLMETLMY (SEQ ID NO: 97)start position 669

158P1D7 Variant 4:
9-mers
IIHSLMKSILWSKASGRGRREE (SEQ ID NO: 98) start position 674
10-mers
NIIHSLMKSILWSKASGRGRREE (SEQ ID NO: 99) start position 673
15-mers
LILAGNIIHSLMKSILWSKASGRGRREE (SEQ ID NO: 100) start position 668

158P1D7 Variant 6:
9-mers
GNIIHSLMNPSFGPKHLEEEEER (SEQ ID NO: 101) start position 372
10-mers
AGNIIHSLMNPSFGPKHLEEEEER (SEQ ID NO: 102) start position 371
15-mers
RKLILAGNIIHSLMNPSFGPKHLEEEEER (SEQ ID NO: 103) start position 366

Table LVI: Protein Characteristics of 158P1D7

	Bioinformatic Program	URL	Outcome
ORF	ORF finder		2555 bp
Protein length			841 aa
Transmembrane region	TM Pred	http://www.ch.embnet.org/	One TM, aa609-aa633
	HMMTop	http://www.enzim.hu/hmmtop/	One TM, aa609-aa633
	Sosui	http://www.genome.ad.jp/SOSui/	One TM, aa608-aa630
	TMHMM	http://www.cbs.dtu.dk/services/TMHMM	One TM, aa611-aa633
Signal Peptide	Signal P	http://www.cbs.dtu.dk/services/SignalP/	Signal peptide, aa3-aa25
pI	pI/MW tool	http://www.expasy.ch/tools/	pI 6.07
Molecular weight	pI/MW tool	http://www.expasy.ch/tools/	95.1 kD
Localization	PSORT	http://psort.nibb.ac.jp/	Plasma membrane
	PSORT II	http://psort.nibb.ac.jp/	65% nuclear, 8% cytoplasmic, 4% plasma membrane
Motifs	Pfam	http://www.sanger.ac.uk/Pfam/	Leucine-rich repeat; mannosyl transferase
	Prints	http://bioinf.man.ac.uk/cgi-bin/dbbrowser	Leucine-rich repeats; Relaxin receptor
	Blocks	http://www.blocks.fhcrc.org/	Leucine rich repeats; cysteine-rich flanking region

Table LVII. Characteristics of 158P1D7 specific antibodies

mAb	Isotype	Affinity (nM)	FACS	Internalization	Western
X68(2)22.1.1	IgG2b/k	3.8	+	+	+
X68(2)31.1.1	IgG2a/k	14	+	+	+
X68(2)18.1.1	IgG2a/k	19	+	+	+
X68(2)120.1.1	IgG2a/k	19	+	+	+

Table LVIII: Detection of 158P1D7 protein by immunohistochemistry in various cancer patient specimens.

TISSUE	Number Positive	Number tested	% No. Positive
Bladder TCC	35	71	49.3
Lung Carcinoma	26	6	23.1
Breast Carcinoma	11	10	90.9